

```

1 TCCTCCGGTC GCCCGCCCTC GGGGCAGCTA GTGGCGCAGC CCCCCGCCCCG
51 CGGCCCCGGC CTCCCCGGCG GCGCGGCAGG GGAGGGGTTA AGCTGCCGCA
101 GGGACCGCCG CGTGCGGGGC GAGAGGGAGC CCCCAGTGGG GGTGGCGCAG
151 CCGGCGGGGT TCGGTCCGAG CCCGGTGGGA GGCTCCCGGA GCGCAGCCTG
201 GGCCCAAGCC ACCCCGCGCC GCGGGCCATG GCAGGCACCC TGGACCTGGA
251 CAAGGGCTGC ACGGTGGAGG AGCTGCTCCG CGGGTGCATC GAAGCCTTCG
301 ATGACTCCGG GAAGGTGCGG GACCCGCAGC TGGTGCGCAT ATTCTCATG
351 ATGCACCCCT GGTACATCCC CTCCTCTCAG CTGGCGGCCA AGCTGCTCCA
401 CATCTACCAA CAATCCCGGA AGGACAACTC CAATTCCTG CAGGTGAAAA
451 CGTGCCACCT GGTTCAGTAC TGGATCTCCG CCTTCCAGC GGAGTTTGAC
501 TTGAACCCCG AGTTGGCTGA GCAGATCAAG GAGTGGAAGG CTCTGCTAGA
551 CCAAGAAGGG AACCAGCGGC ACAGCAGCCT AATCGACATA GACAGCGTCC
601 CTACCTACAA GTGGAAGCGG CAGGTGACTC AGCGGAACCC TGTGGGACAG
651 AAAAAGCGCA AGATGTCCCT GTTGTTTGAC CACCTGGAGC CCATGGAGCT
701 GCGCGAGCAT CTCACCTACT TGGAGTATCG CTCCTTCTGC AAGATCCTGT
751 TTCAGGACTA TCACAGTTTC GTGACTCATG GCTGCACTGT GGACAACCCC
801 GTCCTGGAGC GGTTCATCTC CCTCTTCAAC AGCGTCTCAC AGTGGGTGCA
851 GCTCATGATC CTCAGCAAAC CCACAGCCCC GCAGCGGGCC CTGGTCATCA
901 CACACTTTGT CCACGTGGCG GAGAAGCTGC TACAGTGCA GAACTTCAAC
951 ACGCTGATGG CAGTGGTGGG GGGCCTGAGC CACAGCTCCA TCTCCGCTCCT
1001 CAAGGAGACC CACAGCCACG TTAGCCCTGA GACCATCAAG CTCTGGGAGG
1051 GTCTCACGGA ACTAGTGACG GCGACAGGCA ACTATGGCAA CTACCGGCGT
1101 CGGCTGGCAG CTTGTGTGGG CTTCCGCTTC CCGATCCTGG GTGTGCACCT
1151 CAAGGACCTG GTGGCCCTGC AGCTGGCACT GCCTGACTGG CTGGACCCAG
1201 CCCGACCCG GCTCAACGGG GCCAAGATGA AGCAGCTCTT TAGCATCCTG
1251 GAGGAGCTGG CCATGGTGAC CAGCCTGCGG CCACCAGTAC AGGCCAACCC
1301 CGACCTGCTG AGCCTGCTCA CCGTGTCTCT GGATCAGTAT CAGACGGAGG
1351 ATGAGCTGTA CCAGCTGTCC CTGCAGCGGG AGCCGCGCTC CAAGTCCTCG
1401 CCAACCAGCC CCACGAGTTG CACCCACCA CCCCAGGCCC CGGTACTGGA
1451 GGAGTGGACC TCGGTGCGCA AACCAGAGCT GGATCAGGCC CTCGTGGTGG
1501 AGCACATCGA GAAGATGGTG GAGTCTGTGT TCCGGAACCT TGACGTCGAT
1551 GGGGATGGCC ACATCTCACA GGAAGAATTC CAGATCATCC GTGGGAACCT
1601 CCCTTACCTC AGCGCCTTTG GGGACCTCGA CCAGAACCAAG GATGGCTGCA
1651 TCAGCAGGGA GGAGATGGTT TCCTATTTCG TGCGCTCCAG CTCTGTGTTG
1701 GGGGGGCGCA TGGGCTTCGT ACACAACTTC CAGGAGAGCA ACTCCTTGCG
1751 CCCCCTCGCC TGCCGCCACT GCAAAGCCCT GATCCTGGGC ATCTACAAGC
1801 AGGGCCTCAA ATGCCGAGCC TGTGGAGTGA ACTGCCACAA GCAGTGCAAG
1851 GATCGCCTGT CAGTTGAGTG TCGGCGCAGG GCCCAGAGTG TGAGCCTGGA
1901 GGGGTCTGCA CCCTCACCTC CACCCATGCA CAGCCACCAT CACCGCGCCT
1951 TCAGCTTCTC TCTGCCCCGC CCTGGCAGGC GAGGCTCCAG GCCTCCAGCA
2001 ATCCCCCTCC CAGCAGAGAT CCGTGAGGAG GAGGTACAGA CCGTGGAGGA
2051 TGGGGTGTTC GACATCCACT TGTAATAGAT GCTGTGGTTG GATCAAGGAC
2101 TCATTCTGTC CTTGGAGAAA ATACTTCAAC CAGAGCAGGG AGCCTGGGGG
2151 TGTCGGGGCA GGAGGCTGGG GATGGGGGTG GGATATGAGG GTGGCATGCA
2201 GCTGAGGGCA GGGCCAGGGC TGGTGTCCCT AAGGTTGTAC AGACTCTTGT
2251 GAATATTTGT ATTTTCCAGA TGGAATAAAA AGGCCCGTGT AATTAAAAAA
2301 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA (SEQ ID NO:1)

```

**FEATURES:**

5'UTR: 1-227  
Start Codon: 228  
Stop Codon: 2073  
3'UTR: 2076

**Homologous proteins:****Top 10 BLAST Hits**

	Score	E
CRA 1000682340958 /altid=gi 6358505 /def=gb AAF07219.1 AF043722...	1293	0.0
CRA 18000005086608 /altid=gi 5031623 /def=ref NP_005816.1  RAS ...	1241	0.0
CRA 18000005188697 /altid=gi 6755290 /def=ref NP_035372.1  RAS,...	1202	0.0
CRA 18000005205935 /altid=gi 7662334 /def=ref NP_056191.1  KIAA...	618	e-175
CRA 18000005188699 /altid=gi 3928857 /def=gb AAC79700.1  (AF081...	533	e-150
CRA 18000005152782 /altid=gi 9507035 /def=ref NP_062084.1  RAS ...	531	e-149
CRA 18000005192860 /altid=gi 7242201 /def=ref NP_035376.1  RAS ...	529	e-149
CRA 18000005192861 /altid=gi 4038292 /def=gb AAC97349.1  (AF106...	526	e-148
CRA 18000005188698 /altid=gi 5032025 /def=ref NP_005730.1  RAS ...	525	e-148
CRA 1000733831533 /altid=gi 6650545 /def=gb AAF21898.1 AF081197...	525	e-148

**BLAST dbEST hits:**

	Score	E
gi 5432583 /dataset=dbest /taxon=9606 ...	1310	0.0
gi 9876673 /dataset=dbest /taxon=960...	1281	0.0
gi 11286864 /dataset=dbest /taxon=96...	1249	0.0
gi 11285315 /dataset=dbest /taxon=96...	1207	0.0
gi 5432584 /dataset=dbest /taxon=9606 ...	733	0.0
gi 4372300 /dataset=dbest /taxon=9606 ...	720	0.0
gi 12295751 /dataset=dbest /taxon=96...	700	0.0
gi 12288965 /dataset=dbest /taxon=96...	599	e-168
gi 6920402 /dataset=dbest /taxon=960...	573	e-161
gi 2005039 /dataset=dbest /taxon=9606 ...	573	e-161

**EXPRESSION INFORMATION FOR MODULATORY USE:****library source:****From BLAST dbEST hits:**

gi|5432583 Testis  
gi|9876673 Liver-non-cancerous  
gi|11286864 Brain glioblastoma  
gi|11285315 Brain glioblastoma  
gi|5432584 Testis  
gi|4372300 B Cell Chronic lymphatic leukemia  
gi|12295751 Adult marrow  
gi|12288965 Adult marrow  
gi|6920402 Lymph germinal center B cell  
gi|2005039 Lymph

**From tissue screening panels:**

Leukocyte

```

1 MAGTLDLDKG CTVEELLRGC IEAFDDSGKV RDPQLVRIFL MMHPWYIPSS
51 QLAAKLLHIY QQSRKDNSNS LQVKTCHLVR YWISAFPAEF DLNPELAEQI
101 KELKALLDQE GNRHRSSLID IDSVPTYKWK RQVTQRNPVG QKKRKMSLLF
151 DHLEPMELAE HLTYLEYRSF CKILFQDYHS FVTHGCTVDN PVLERFISLF
201 NSVSQWVQLM ILSKPTAPQR ALVITHFVHV AEKLLQLQNF NTLMAVVGGL
251 SHSSISRLKE THSHVSPETI KLWEGLTELV TATGNYGNYR RRLAACVGFR
301 FPILGVHLKD LVALQLALPD WLDPARTRLN GAKMKQLFSI LEELAMVTSL
351 RPPVQANPDL LSLTLVSLDQ YQTEDELYQL SLQREPRSKS SPTSPTSCTP
401 PPRPPVLEEW TSAAKPKLDQ ALVVEHIEKM VESVFRNFDV DGDGHISQEE
451 FQIIRGNFPY LSAFGDLQDN QDGCISREEM VSYFLRSSSV LGGRMGFVHN
501 FQESNSLRPV ACRHCKALIL GIYKQGLKCR ACGVNCHKQC KDRLSVECRR
551 RAQSVSLEGS APSPSPMHSR HHRAFSFSLP RPGRGRSRPP AIPLPAEIRE
601 EEVQTVEDGV FDIHL (SEQ ID NO:2)

```

#### FEATURES:

##### Functional domains and key regions:

[1] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE

cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

```

1      113-116 RRHS
2      144-147 RKMS
3      584-587 RRGs

```

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE

Protein kinase C phosphorylation site

Number of matches: 7

```

1      27-29 SGK
2      63-65 SRK
3      126-128 TYK
4      134-136 TQR
5      269-271 TIK
6      349-351 SLR
7      506-508 SLR

```

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE

Casein kinase II phosphorylation site

Number of matches: 9

```

1      12-15 TVEE
2      63-66 SRKD
3      117-120 SLID
4      163-166 TYLE
5      339-342 SILE
6      373-376 TEDE
7      447-450 SQEE
8      476-479 SREE
9      605-608 TVED

```

[4] PDOC00008 PS00008 MYRISTYL

N-myristoylation site

Number of matches: 4

```

1      19-24 GCIEAF
2      249-254 GLSHSS
3      284-289 GNYGNY
4      492-497 GGRMGF

```

[5] PDOC00009 PS00009 AMIDATION  
Amidation site

582-585 PGRR

[6] PDOC00018 PS00018 EF\_HAND  
EF-hand calcium-binding domain

Number of matches: 2

1 439-451 DVDGDGHISQEEF  
2 468-480 DQNQDGCISREEM

[7] PDOC00379 PS00479 DAG\_PE\_BIND\_DOM\_1  
Phorbol esters / diacylglycerol binding domain

499-548 HNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVEC

**Membrane spanning structure and domains:**

Helix	Begin	End	Score	Certainty
1	34	54	0.713	Putative
2	195	215	0.653	Putative
3	238	258	0.788	Putative

**BLAST Alignment to Top Hit:**

>CRA|18000005086608 /altid=gi|5031623 /def=ref|NP\_005816.1| RAS guanyl  
releasing protein 2 (calcium and DAG-regulated); calcium  
and diacylglycerol-regulated guanine nucleotide exchange  
factor I [Homo sapiens] /org=Homo sapiens /taxon=9606  
/dataset=nraa /length=609

Length = 609

Score = 1241 bits (3176), Expect = 0.0

Identities = 608/615 (98%), Positives = 609/615 (98%)

Frame = +3

Query: 228 MAGTLDLDKGCTVEELLRGCTAEFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIY 407  
MAGTLDLDKGCTVEELLRGCTAEFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIY  
Sbjct: 1 MAGTLDLDKGCTVEELLRGCTAEFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIY 60

Query: 408 QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 587  
QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID  
Sbjct: 61 QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 120

Query: 588 IDSVPTYKWKQVQTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 767  
IDSVPTYKWKQVQTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS  
Sbjct: 121 IDSVPTYKWKQVQTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 180

Query: 768 FVTHGCTVDNPNVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 947  
FVTHGCTVDNPNVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF  
Sbjct: 181 FVTHGCTVDNPNVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 240

Query: 948 NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNRRRLAACVGFR 1127  
NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNRRRLAACVGFR  
Sbjct: 241 NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNRRRLAACVGFR 300

Query: 1128 FPILGVHLKDLVALQALPDWLDPARTRNGAKMKQLFSILEELAMVTSLRPPVQANPDL 1307  
FPILGVHLKDLVALQALPDWLDPARTRNGAKMKQLFSILEELAMVTSLRPPVQANPDL  
Sbjct: 301 FPILGVHLKDLVALQALPDWLDPARTRNGAKMKQLFSILEELAMVTSLRPPVQANPDL 360

Query: 1308 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 1487  
LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ  
Sbjct: 361 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 420

Query: 1488 ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 1667  
ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM  
Sbjct: 421 ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 480

Query: 1668 VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847  
VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC  
Sbjct: 481 VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540

Query: 1848 KDRLSVECRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRGSRPPAIPAEIRE 2027  
KDRLSVECRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRGSRPP EIRE  
Sbjct: 541 KDRLSVECRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRGSRPP-----EIRE 594

Query: 2028 EEVQTVEDGVFDIHL 2072  
EEVQTVEDGVFDIHL  
Sbjct: 595 EEVQTVEDGVFDIHL 609 (SEQ ID NO:4)

>CRA|1000682340958 /altid=gi|6358505 /def=gb|AAF07219.1|AF043722\_1  
 (AF043722) guanine exchange factor MCG7 isoform 1 [Homo  
 sapiens] /org=Homo sapiens /taxon=9606 /dataset=nraa  
 /length=671  
 Length = 671  
 Score = 1293 bits (3309), Expect = 0.0  
 Identities = 639/670 (95%), Positives = 643/670 (95%), Gaps = 4/670 (0%)  
 Frame = +3

Query: 75 GRGGVKLPQGPFRAGREGAPGGGGAAG---GVRSEPGGRLPERSLGAHPAPAAMAGTL 242  
 GRG P + +E G +G GVRSEPGGRLPERSLGAHPAPAAMAGTL  
 Sbjct: 8 GRGTQGWPGSSEQHVQEATSSAGLHSGVDELGVRSEPGGRLPERSLGAHPAPAAMAGTL 67

Query: 243 DLDKGCTVEELLRGCI EAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRK 422  
 DLDKGCTVEELLRGCI EAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIYQQSRK  
 Sbjct: 68 DLDKGCTVEELLRGCI EAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIYQQSRK 127

Query: 423 DNSNSLQVKTCHLVRYWISAFPAEFDLPNELAEQIKELKALLDQEGNRRHSSLIDIDSVP 602  
 DNSNSLQVKTCHLVRYWISAFPAEFDLPNELAEQIKELKALLDQEGNRRHSSLIDIDSVP  
 Sbjct: 128 DNSNSLQVKTCHLVRYWISAFPAEFDLPNELAEQIKELKALLDQEGNRRHSSLIDIDSVP 187

Query: 603 TYKWKQVQTQRNPFVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 782  
 TYKWKQVQTQRNPFVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG  
 Sbjct: 188 TYKWKQVQTQRNPFVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 247

Query: 783 CTVDNPVLERFISLFNSVSQWVQMLILSKPTAPQRALVITHFVHVAEKLQLQNFNTLMA 962  
 CTVDNPVLERFISLFNSVSQWVQMLILSKPTAPQRALVITHFVHVAEKLQLQNFNTLMA  
 Sbjct: 248 CTVDNPVLERFISLFNSVSQWVQMLILSKPTAPQRALVITHFVHVAEKLQLQNFNTLMA 307

Query: 963 VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRRPILG 1142  
 VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRRPILG  
 Sbjct: 308 VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRRPILG 367

Query: 1143 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLT 1322  
 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLT  
 Sbjct: 368 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLT 427

Query: 1323 VSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPRPPVLEEWTSAAKPKLDQALVVE 1502  
 VSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPRPPVLEEWTSAAKPKLDQALVVE  
 Sbjct: 428 VSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPRPPVLEEWTSAAKPKLDQALVVE 487

Query: 1503 HIEKMVESVFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLQNDGCISREEMVSYFL 1682  
 HIEKMVESVFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLQNDGCISREEMVSYFL  
 Sbjct: 488 HIEKMVESVFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLQNDGCISREEMVSYFL 547

Query: 1683 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS 1862  
 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS  
 Sbjct: 548 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS 607

Query: 1863 VECRRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRGSRPPAIPPAEIREEEVQT 2042  
 VECRRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRGSRPP EIREEEVQT  
 Sbjct: 608 VECRRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRGSRPP-----EIREEEVQT 661

Query: 2043 VEDGVFDIHL 2072  
 VEDGVFDIHL  
 Sbjct: 662 VEDGVFDIHL 671 (SEQ ID NO:5)

```

>CRA|18000005188697 /altid=gi|6755290 /def=ref|NP_035372.1| RAS,
    guanyl releasing protein 2; RAP 1A protein-specific
    guanine nucleotide exchange factor 1; CalDAG-GEFI [Mus
    musculus] /org=Mus musculus /taxon=10090 /dataset=nraa
    /length=608
    Length = 608
    Score = 1202 bits (3076), Expect = 0.0
    Identities = 589/615 (96%), Positives = 597/615 (96%)
    Frame = +3

Query: 228  MAGTLDLDKGCTVEELLRGCEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIY 407
           MA TLDLDKGCTVEELLRGCEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLA+KLLH Y
Sbjct: 1    MASTLDLDKGCTVEELLRGCEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLASKLLHFY 60

Query: 408  QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 587
           QQSRKDNSNSLQVKTCHLVRYW+SAFPAEFDLNPELAE IKELKALLDQEGNRRHSSLID
Sbjct: 61   QQSRKDNSNSLQVKTCHLVRYWVSAFPAEFDLNPELAEPIKELKALLDQEGNRRHSSLID 120

Query: 588  IDSVPTYKWKQVQTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 767
           I+SVPTYKWKQVQTQRNPV QKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS
Sbjct: 121  IESVPTYKWKQVQTQRNPVEQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 180

Query: 768  FVTHGCTVDNPNVLERFISLNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 947
           FVTHGCTVDNPNVLERFISLNSVSQWVQLMILSKPTA QRALVITHFVHVAEKLLQLQNF
Sbjct: 181  FVTHGCTVDNPNVLERFISLNSVSQWVQLMILSKPTATQRALVITHFVHVAEKLLQLQNF 240

Query: 948  NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNRRRLAACVGFR 1127
           NTLMAVVGGLSHSSISRLKETHSHVSP+TIKLWEGLTELVTATGNY NYRRRLAACVGFR
Sbjct: 241  NTLMAVVGGLSHSSISRLKETHSHVSPDTIKLWEGLTELVTATGNYSNYRRRLAACVGFR 300

Query: 1128  FPILGVHLKDLVALQALPDWLDPARTRNGAKMKQLFSILEELAMVTSLRPPVQANPDL 1307
            FPILGVHLKDLVALQALPDWLDP RTRNGAKM+QLFSILEELAMVTSLRPPVQANPDL
Sbjct: 301  FPILGVHLKDLVALQALPDWLDPGRTRNGAKMRQLFSILEELAMVTSLRPPVQANPDL 360

Query: 1308  LSLTLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPRPPVLEEWTSAAPKLDQ 1487
            LSLTLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPRPPVLEEWTSAAPKLDQ
Sbjct: 361  LSLTLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPRPPVLEEWTSAAPKLDQ 420

Query: 1488  ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEM 1667
            ALV EHEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEM
Sbjct: 421  ALVAEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEM 480

Query: 1668  VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
            +SYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC
Sbjct: 481  ISYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540

Query: 1848  KDRLSVECRRRRAQSVSLEGSAPSPSPMHHHRAFSFSLPRPGRGSRPPAIPAEIRE 2027
            KDRLSVECRRRRAQSVSLEGSAPSP H+ HHRAFSFSLPRPGRR SRPP EIRE
Sbjct: 541  KDRLSVECRRRRAQSVSLEGSAPSPSPTHT-HHRAFSFSLPRPGRSSRPP-----EIRE 593

Query: 2028  EEVQTVEDGVFDIHL 2072
            EEVQTVEDGVFDIHL
Sbjct: 594  EEVQTVEDGVFDIHL 608 (SEQ ID NO:6)

```

>CRA|18000005205935 /altid=gi|7662334 /def=ref|NP\_056191.1| KIAA0846  
protein [Homo sapiens] /org=Homo sapiens /taxon=9606  
/dataset=nraa /length=689  
Length = 689  
Score = 618 bits (1576), Expect = e-175  
Identities = 314/597 (52%), Positives = 409/597 (67%), Gaps = 5/597 (0%)  
Frame = +3

Query: 234 GTLDLDKGCTVEELLRGCI EAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQ 413  
G+ L K T++ELL CIE FDD+G++ + L RI L+MH WY+ S++LA KLL +Y+  
Sbjct: 2 GSSGLGKAATLDELLCTCIEMFDDNGELDNSYLPRIVLLMHRWYLSSTELAEKLLCMYRN 61

Query: 414 SRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDID 593  
+ ++ N ++K C+ +RYWI FPAEF+L+ L +E + + Q G +H SLIDI  
Sbjct: 62 ATGESCFNEFRLKICYFMRYWILKFPAEFNLDLGLIRMTEEFREVASQLGYEKHVSLLIDIS 121

Query: 594 SVPTYKWKQRQVTQRNPVGQKKRMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFV 773  
S+P+Y W R+VTQR V KK K LLFDHLEP+ELAEHLT+LE++SF +I F DY S+V  
Sbjct: 122 SIPSYDWMRRVTQRKKVS-KKGACLLFDHLEPIELAEHLTFLEHKSFRIRISFTDYQSYV 180

Query: 774 THGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFT 953  
HGC +NP LER I+LFN +S+WWQLM+LSKPT QRA VIT F++VA+KLLQL+NFT  
Sbjct: 181 IHGCLENNPTLERSIALFNGISKWVQLMVLSKPTPQQRAEVITKFINVAKLLQLKNFT 240

Query: 954 LMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFP 1133  
LMAVVGGLSHSSISRLKETHSH+S E K W +TELV++ GNY NYR+ A C GF+ P  
Sbjct: 241 LMAVVGGLSHSSISRLKETHSHLSSEVTKNWNEMTELVSSNGNYCNRYKAFADCDGFKIP 300

Query: 1134 ILGVHLKDLVALQALPDWLDPARTRLNAGKMKQLFSILEELAMVTSLRPPVQANPDLLS 1313  
ILGVHLKDL+A+ + PDW + ++N KM QL L EL + + ++ N DL++  
Sbjct: 301 ILGVHLKDLIAVHVIFPDWTE--ENKVNIVKMHQLSVTSELVSLQNASHHLEPNMDLIN 358

Query: 1314 LLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQAL 1493  
LLT+SLD Y TED++Y+LSL EPR+ SPTSPT+ P +P V EW PK D +  
Sbjct: 359 LLTSLDLYHTEDDIYKLSLVLEPRNSKSPTSPTT---PNKPVVPLEWALGVMPKPDPTV 415

Query: 1494 VVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLQNDGCISREEMVS 1673  
+ +HI K+VESVFRN+D D DG+ISQE+F+ I NFP+L +F LD++QDG IS++EM++  
Sbjct: 416 INKHIRKLVESVFRNYDHDHDGYISQEDFESIAANFPFLDSFCVLDKQDGLISKDEMMMA 475

Query: 1674 YFLRSSSVLGGRM--GFVHNQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847  
YFLR+ S L +M GF+HNFQE L+P C HC + GI KQG KC+ CG NCHKQC  
Sbjct: 476 YFLRAKSQLHCKMGPFIHNFQEMTYLKPTFCEHCAGFLWGIKQGYKCKDCGANCHKQC 535

Query: 1848 KDRLSVECRRAQSVSL---EGSAPSPSPMHSHHRAFSFSLPRPGRGRSRPPAIP 2009  
KD L + CRR A++ SL GS P + F F G R AI L  
Sbjct: 536 KDLLVLACRRFARAPSLSSGHGSLPGSPSLPPAQDEVFEFPGVTAGHRDLDSRAITL 592 (SEQ  
ID NO:7)

>CRA|18000005188699 /altid=gi|3928857 /def=gb|AAC79700.1| (AF081196)  
calcium and DAG-regulated guanine nucleotide exchange  
factor II [Rattus norvegicus] /org=Rattus norvegicus  
/taxon=10116 /dataset=nraa /length=795  
Length = 795  
Score = 533 bits (1358), Expect = e-150  
Identities = 267/590 (45%), Positives = 390/590 (65%), Gaps = 12/590 (2%)  
Frame = +3

Query: 156 GVRSEPGGRLPERSLGPAPAPAMAGTLD-----LDKGCTVEELLRGCI EAFDD 308  
G R+ P GRL +S PA ++A L KG +++L+ CI++FD  
Sbjct: 17 GSRAGPKGRLEAKSTNSPLPAQPSLAQITQFRMMVSLGHLAKGASLDDLIDSCIQSFDAD 76

FIGURE 2, page 6 of 7



Query: 309 GKV-RDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRKDNSNSLQVKTCHLVRYWISAF 485  
 G + R QL+++ L MH I S++L KL+++Y+ + + NS + +K C+ VRYWI+ F  
 Sbjct: 77 GNLCRSNQLLQVMLTMHRIIISAEALLQKLMNLYKDALEKNSPGICLKICYFVRYWITEF 136

Query: 486 PAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVPTYKWKRQVTQRNPVG-QKKRK 662  
 F ++ L ++E + L+ G H LID + + W R++TQR KKRK  
 Sbjct: 137 WIMFKMDASLTSTMEEFQDLVKANGEESHCHLIDTTQINSRDWSRKLQRIKSNTSKKRK 196

Query: 663 MSLFLDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHGCTVDNPVLERFISLFNSVSQ 842  
 +SLFLDHLEP EL+EHLYTYLE++SF +I F DY +++ + C +NP +ER I+L N +SQ  
 Sbjct: 197 VSLFLDHLEPEELSEHLTYLEFKSFRRISFSQYQNYLVNSCVKENPTMERSIALCNGISQ 256

Query: 843 WVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMAVVGGLSHSSISRLKETHSHV 1022  
 WVQLM+LS+PT RA V F+HVA+KL QLQNFNTLMAV+GGL HSSISRLKET SHV  
 Sbjct: 257 WVQLMVLRSRTPQLRAEVFIKFIHVAQKLHQLQNFNTLMAVIGGLCHSSISRLKETSSHV 316

Query: 1023 SPETIKLWEGLTELVTATGNYGNYYRRRLAACVGFRRFPILGVHLKDLVALQLALPDWLDPA 1202  
 E K+ +TEL+++ NY NYRR C F+ PILGVHLKDL++L A+PD+L+  
 Sbjct: 317 PHEINKVLGEMTELLSSCRNYDNYRRAYGECTHFKIPILGVHLKDLISLYEAMPDYLEDG 376

Query: 1203 RTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLTVSLDQYQTEDELYQLSLQRE 1382  
 ++N K+ L++ + EL + + PP+ AN DL+ LLT+SLD Y TEDE+Y+LS RE  
 Sbjct: 377 --KVVNQKLLALYNHINELVQLQDVAPPLDANKDLVHLLTSLDLYYTEDEIYELSYARE 434

Query: 1383 PRSKSSPTSPTSCTPPPRPPVLEEWTSAAPKLDQALVVEHIEKMVESVFRNFDVDGDGH 1562  
 PR+ +P P +PPV+ +W S PK D + +H+++MV+SVF+N+D+D DG+  
 Sbjct: 435 PRNHRAPP-----LTPSKPPVVVDWASGVSPKDPKTIKSHVQRMVDSVFKNYDLQDGY 489

Query: 1563 ISQEEFQIIRGNFPYLSAFGDLQNDGDCISREEMVSYFLRSSSVLGG-RMGFVHNFQES 1739  
 ISQEEF+ I +FP+ +F +D++++G ISR+E+ +YF+R+SS+ +GF HNFQE+  
 Sbjct: 490 ISQEEFEKIAASFPP--SFCVMDKDRGLISRDEITAYFMRASSIYSKLGGLGFPHNFQET 547

Query: 1740 NSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVECCRRAQS 1889  
 L+P C +C + G+ KQG +C+ CG+NCHKQCKD + EC++R++S  
 Sbjct: 548 TYLKPTFCDCNAGFLWGVKQGYRCKDCGMNCHKQCKDLVVFCKKRSKS 597 (SEQ ID  
 NO:8)

#### Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00617	RasGEF domain	123.5	4e-33	1
PF00130	Phorbol esters/diacylglycerol binding domain	59.5	3.6e-14	1
PF00036	EF hand	21.8	0.00027	2
PF01237	Oxysterol-binding protein	3.5	4.2	1

#### Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01237	1/1	249	272 ..	1	24 [.	3.5	4.2
PF00617	1/1	148	336 ..	1	227 [.]	123.5	4e-33
PF00036	1/2	430	458 ..	1	29 [.]	17.4	0.0047
PF00036	2/2	463	482 ..	5	24 ..	6.7	4.9
PF00130	1/1	499	548 ..	1	51 [.]	59.5	3.6e-14

1 ACAGAAAGGT CCTGTTTCTA AGTCTTACAT TACCAAGACT GAGGTGCGGG  
51 GGCGGTCTCTG GATCCCCCGC CCCAAGGCTG GGAGGGGCAC GCCTCGGAAG  
101 GGAGGTTTGG GGTCCGGTGGT TTCACAGTGA GTGTGTCTGA AGCCAAATGG  
151 TCGGAAACCG TTACCCGCTC TCCTAGGCCC GGCTAGTGGG GACCCCAACC  
201 GCCTGCGGCT GCCCCCTCCA AGTTCTCTCC GTTGGCCAG GCATCCAGGT  
251 CTCCAGTCTC CGAGCTGCGG AGAACCCACC GCCACATGCG GCTGCCCTT  
301 TCCATTTCGAC CCTGTGGGGA GCCAGGCTTC CGGGGCCCCG TTCCTCCTGT  
351 GTGAACTGGG CCCCCCGCCC CCATTCCCAG ACATCAAGGC CGCTCTCCA  
401 GATAGCCACG ATTTTCATTCC TCGTCCCCA CAGGTCCCTC TCCCCAAAT  
451 ATTCCCATCT TGTCCTAGCC CATCCCCCAG ACTATCTCAA GGACCAGCTG  
501 TCCCCACGCC CCCGACCTCC ACTAGGCCTG TGCCACCCGC TGCCTGCAGG  
551 AAGACGCCCCG GTCCCGGGCC GGGTTAGCCC CATGGGAACG GTTTGTCTCG  
601 AAAACAGGAA CCCGAGCTGG GGGCTGGGCG GGGCGCCCCT TCCCCACCGC  
651 AGTCCGCTTC CTGCCCCCTC CGGCTTCCTC CGCCCGACAC CCAGGCAGGG  
701 CGGGGGGCAC TGGGGCGTCC GCGGTTGGGG GAGGGGCTCT TCGTTTCGGT  
751 CCCCCCTCCC GCGTCCCGGG CGGCGGGGCC TCCGGTCGCC CGCCTCGGGG  
801 CAGCTAGTGG CGCAGCCCCC CGCCCGCGGC CCTGGCCTCC CGGGCGGCGC  
851 GGCAGGGGAG GGGTTAAGCT GCCGCAGGGA CCGCCGCGTG CGGGGCGAGA  
901 GGGAGCCCCC GGTGGGGGTG GCGCAGCCGG CGGTGCGGAG CTCCGCGCAG  
951 GGGCGGAGGG GGGAGGGGGC AGCCTGGCGC GGGGGCGGGG GCGGGGCGGC  
1001 GGGGAGCGGG GCCGCGGCGT GGAGAGCGGG CGGGAGCCGC AGCCGCAGCG  
1051 AGGCCGCGGG GCGGGAGCGC ACGGAGGTGG GGTCCGCCAG GCCGGTGGCG  
1101 GTCCTTGCG GCAGGTCCCA AGAGTGAGTG GGCAGCGCG GCGGGGCGC  
1151 CAGGCGAAGG AGGGCGCGGC CCCCAGCGAC TCCCCCCCCG CCCAGGGCGG  
1201 CGCGGGCGGG CTGGGGGCGG CGAGCGGGTG GGGAGTCTGC GGCCCGGGT  
1251 TGGGAGAGGG GGCAGCGGCC ACGAGAGCTA AGGCGCGCTG GATCCCCGGA  
1301 GGGCGGAGGA CCTCCACGGT GCACCCAGCT TTTCCAGCC ACCTTCCAGC  
1351 GGGGCCCTCC CCCGCGTACC CCCATTGGG AGATGAGAAA ATTGAGGCTC  
1401 CCAGAGGCCA AGTGATTCTC AAGGTCACAC GAGGAAGCGG TAGAGCCAGG  
1451 CGGGGACGGC TCTGGGTGGC TCTTAGGAAA AGTCCGCCTG AGAACTCCGT  
1501 ACAGGAGCTC CCTGTCTCTC CAGCCTGGGG GAGTGAGTAT GTGTAGGGCC  
1551 GGGGTACCTT TCCGTGGGGC AAGGCTCTGC CAAAATCTGG GAGTGAGGGG  
1601 AGTCAGGGAG CTGGGGCCGC AGGGCGGGCC CTGCACCGCA AATGGGAGGG  
1651 GGGCGACGGA ATGGGCGTGC GCACCCATGG GGGTGTGTGC ATGTGTGTGG  
1701 GAGTGTACAT GCGTGGAGAG GCACTGCCTT GCGTGTGTGC ACACGTGTGA  
1751 GGATGTCAGC GCCTGTGTGG CCGCGGGACT CAAGGCTGGC CTGGCTCAAG  
1801 TGAACAGCAC GTCCAGGAGG CGACCTCGTC CGCGGGTTTG CATCTGGGG  
1851 TGGACGAGCT GGGTATGTGT GCCTGAGGGT TTCTTCGTGC AGGTGTGCAC  
1901 AGGGTGTGGG TGCCATTGTG TGTGAGAGAC GGAGGATGGG GAGGCCGGTG  
1951 CCTGTGGCCC GGTGCGTGTA AGTGCGGACG CCTGCACCTC CACTTAGGTC  
2001 CCCGCCCTCC GACGACTAAC TTGGGTGTGG AGTGTTTGCC CCTGCCAGGG  
2051 TGCGTATGAC CCCGCCAGTG ACCGGAGTTG CTAATGGTGT CATGCACCCA  
2101 CCGGCCACCC TTGGCGCGAG CGCCCCCTC TGGACACCCT GCTCCGTGCG  
2151 CGCTCACAGT TCGCCTGTGC GGGGCCGGGG CCAGGGTCAG GAGCCGGGGA  
2201 TAGGGAGGAA GAGGGCCTGT GGACAAGCTG AGCCGGGACC CCTGGGACCT  
2251 TTGCGGAGGT GGCCTGGGAG CGCTCAGTTC CCAGGCTGAG GCTTCCCGCT  
2301 GACGCCCTCT GGCCGCAGCG GGCTCCCCC GCCCAGGAA TGTTCTCTC  
2351 CCATCCAGTC CGCCTCCCCT AGGGCAGGCC CCCTGGGGGC TGCCGAGCC  
2401 CCGCTCGCC TTCTTGGGCT CCCGGGAGGG GCGGAGGCGA GCAGGACGCC  
2451 TGGGTTCTCT CCCCCACCT CCCATACCAG GGAGAAATTC CTCCGAGGTC  
2501 CCCTCAGGCT CTGGGTTCCC AAAATAACCC TGCGGGGGAA GGGAGGCTGT  
2551 GGAGGGAGGG AAGCGGGAGG GGCGCAGAGC CGAGCTGCGG GGTGCTGCAG  
2601 GTGCCTCTGG GGAGAGGGCG CGAGGAGAAG GCGCCCTGCG GGGGGTGGG  
2651 CGCCAGCCAG TCCTGGGATC TTGGTTCTGC CCCATCCTCG TGAAGCCCCT  
2701 CGGCCTTCCC GCGACTCCGA GGGTGGGCCG GAAGCCTCTC TGCGGGTCCG  
2751 TTTCCCAACT GCGGGGTTGC ACCATCCCGG GCCAGACCGT TTAACCCCGG  
2801 GAGTGGCCGC GGGGACAAC TCCGCCCTG TCCAGCAGGG GCGGTGCCCG  
2851 CCCC GCCCGG TTTCTGCCCC CGGGGCCGCT CCCCCGCCG GCACTCCGCA  
2901 GACTCCCGCT CTGCCTCTCC CGGGACAGGG GTTCGGTCCG AGCCCGGTGG  
2951 GAGGCTCCCG GAGCGCAGCC TGGGCCAGC CCACCCCGCG CCGGCGGCCA  
3001 TGGCAGGCAC CCTGGACCTG GACAAGGGCT GCACGGTGA GGAGCTGCTC  
3051 CGCGGGTGCA TCGAAGCCTT CGGTGAGTGG CTCGGGAGGG CACACGGAGC  
3101 CTGAGCCTAG CCCCAGTCT GAGCCCGGGT CCCTGCCTCC CAGGCACAGT

FIGURE 3, page 1 of 12

3151 CCAGGGCACA GCCCTGACCC GGACCCACCC TGCTCCGCAG CGTGCAGTCT  
3201 CTTTAAACGAA AGCCTCCTCC GCAACGCAGG GCAGAGAGAT GCACGCCCTT  
3251 CAGACAGATG AGGTTTCCCT TCTCTAGCCT TCCCCAGCGG CGGCGAAGGG  
3301 AGGGCCGGGT CCCGGACTCT GACACTTGAG GGGCATTATC TGTCTCCCGG  
3351 GGAATCCGGA GGAACCTCGT ATCTCCGGCC TGGGAGCTGT TTCCGGCTAA  
3401 TGGGGGGCGG CTTATCTGGT GAAGGGGTGC CCCTTCCCC CAAGCGCTCA  
3451 GGAAATGACC TCTGGATTCT TGACCCCGGG GAACCCAGGC TCCTTCCGCC  
3501 CCAGCTGGTT CCCCTCCGGA CGATGGGCGG CTCGGGCGCT CCCCTCCTCC  
3551 AGTCCTCAGG GCGTGCCAT CTCTCGCCCA CCACACCTTT CCTCTCTAAT  
3601 TTGCCTCCTG CTCTCGGAGT CCTGGGCAAG CAGGAGGTGG GCGGGGTCGA  
3651 GCGTGACCC GAAGGACCGA TACCTGGCGG GTTGCGGGGT GAGGATGAGG  
3701 CATGGTAGCT GCGGACCCAG CTCAGCCACC TGTCTTTGAC CCTTCGGAGT  
3751 CAGATGACTC CGGGAAGGTG CGGGACCCGC AGCTGGTGCG CATGTTCCCTC  
3801 ATGATGCACC CCTGGTACAT CCCCTCCTCT CAGCTGGCGG CCAAGCTGCT  
3851 CCACATATAT CTTTCGCCGG CCTTGCCAAG GCCCCGCGG TCGGAGCCCA  
3901 TGCGCAGCCC CTCTGCCAG CCCAGTGCA GAATGAGCCT CGCTCCTAAG  
3951 TATAGGCCAC TCCTTATCCC AGAGCTCAGG CGTCGTCCA GCCTCCAAC  
4001 AGGGCCTAGG CTCTGCCCCC TCCTTGCTCC TAGCGACTCG GTCCTGTCCC  
4051 CAGGCTCTGT CCCCAGCCGA GGCCTTGCCC TCCTTCTCCC TAGAGTCTAG  
4101 GGCCTGCCCC TGCTTCAGGC TTGGGTGCGC CCCGTGCATC TCTCTCTCCC  
4151 AGAGCCCAGG CTTTGCTTTC AGCCTCCCTC AGCACCTAGT CCTCCACCCC  
4201 CACCTCCAAC CCCTCCCAGA GCTCAAGCCT CACCCCCAGC ATCTCCGCAG  
4251 AGCGCAAGCC CCATCCCTAG AACGTGTCTC CTAGAACCAG GCCCCGCCCC  
4301 CAGCCTCCCT CCACGCAGGC CTCCTTTCT AGAGTTAAGC GGCTCCTTA  
4351 ACCCTCTCCT TCACCTACCA ACAATCCCGG AAGGACAACT CCAATTCCCT  
4401 GCAGGTGAAA ACGTGCCACC TGGTCAGGTG AGTCTTTCCC CTGGGGCTCT  
4451 AGCCCCCTCC CTTTCTCCCT TCTCTCTGGC TTCAGGCTGG CCTGGAGGAG  
4501 GGGGCAGGGC GCTGTTTCTG GGAGTGGGTT TGAACCTGG CTGTGCCGGG  
4551 TGGCAGTGTC TGCCACAGGC TCACCCCTTC CTGGGTCTGG GCCTTAATTT  
4601 TCTTTTCTGC GCAGTGCGGG TGGTTGTCTC AAGGGTCTAA TGTACACTTG  
4651 GAGTGCGGAA GGAAAGAGCT GGAACCATAG TTTGAGGGTC TTTTGTCTTA  
4701 GGTGACTATA ATCTCAAATA GCTCCTTGCA GCCTGCTGGG TGATGGTGGG  
4751 GGAAGGGCTA TCTTGGGTGA CTCCCCGCTC CTCCAGGTAC TGGATCTCCG  
4801 CCTTCCCAGC GGAGTTTGAC TTGAACCCGG AGTTGGCTGA GCAGATCAAG  
4851 GAGCTGAAGG CTCTGCTAGA CCAAGAAGGG AACCACGGC ACAGCAGCCT  
4901 AATCGACATA GACAGCGTGT GCGTGGGGGG AGCACAGAG GCTGGGGGGG  
4951 CACTCAGTAT CCTATACCAT CTGTGCTTAA TAAATGTCTG TTGAACTGAA  
5001 TGAGTGAGGG TCATGTTGCT CTCTCGCTTA AAAACCTTCC ATGGCTCCCT  
5051 ATTGCCTTCA ACATGCCTCC TCTGGGCAGC TTGGCGTTCC TGCCTCATCT  
5101 TCCACTGCCA CCACCCATCC CACACACCTC CTCCTGTAGC TGCGCTGGGT  
5151 CGGCTCCCCG TCGCTGAGC TCTCGAGTCC TTTCTCATCA TGGTGTCTG  
5201 CTCATATCAT CCCCCTTGCT GCCTCCTCCG TGTTACCAAG ACTCAGTTCA  
5251 GGCATGAAGT CTCCTGTTGGT TCTGAGGGT CGGGGCTCTT CCGGGGTAGA  
5301 ATTTGTCTGT CCACCTCTG TTTTCCATGG CACTTTGTAC AGACTCCTGT  
5351 ACAAAGACCT CTGTACATGT GTCACGCTGT TTTGTGATCA TGTGTTCTG  
5401 TGTCTGTCTC CCTCAGTAGA CTGTGAGCTC CTCGAGGGCA GGAACCGTGT  
5451 CTTACTCATC TCTGTATTCC CAGCGCCTAG CACAGTGCCT GGCACAGAGT  
5501 ACGTTGTTC TAAATGTGTG TTGAGTGCAT GACGGGGTGG GGGGAGATGA  
5551 GGAGGAGTTG CTGGGACTGG GAACATTCTG GCCTAGGACA GTGCCTCGCA  
5601 TTATGTAGGT TCTCAGTAAG CGTGAATGGT GTGTCTGTGT GAGTGGGGGG  
5651 CCACGAGGCA TGCGCATGTC CAGCAAAGGG CTCACTACCC CTGCCCCCCC  
5701 AGCCCTACCT ACAAGTGGAA GCGGCAGGTG ACTCAGCGGA ACCCTGTGGG  
5751 ACAGAAAAAG CGCAAGATGT CCCTGTTGTT TGACCACCTG GAGCCCATGG  
5801 AGCTGGCGGA GCATCTCACC TACTTGGAGT ATCGCTCCTT CTGCAAGATC  
5851 CTGGTGCGGC CCGAGGGCTG GGGGGTCAGG GGTCCAATGT GGGCTGGAAG  
5901 AGAGTTCTAG GAGGGGCAGG GTCCCTGGCG TAGGCTGGGT CACAGGGTGC  
5951 ATCAGGGGTT TCAGTTAAC CACTGAAGGT CAGCTGGAGG GTGAGGAGTG  
6001 GCTATCAGTG AGGGGAGAGG CCGGCAAGGT GCTGAGGCCA CTCCTCATGC  
6051 CCCCAGTTTC AGGACTATCA CAGTTTCGTG ACTCATGGCT GCACTGTGGA  
6101 CAACCCCGTC CTGGAGCGGT TCATCTCCCT CTTCAACAGC GTCTCACAGT  
6151 GGGTGCAGCT CATGATCCTC AGCAAACCCA CAGCCCCGCA GCGGGCCCTG  
6201 GTCATCACAC ACTTTGTCCA CGTGGCGGAG GTGCCTGCCC CTCCTCCCG  
6251 GTGTCTCCA ACCACCCAC ATGCCAGTCA GGCCAACCCT TCCCTTCCCC

FIGURE 3, page 2 of 12

6301 TAACCCACTG CTTTCTCTCT AGATAAGCTG GGCCAAATTC TGGGCCCCACT  
6351 CAGTGACTCC CTGCCTCTCC GTCCCCATTT GCCTTCCAGA AGCTGCTACA  
6401 GCTGCAGAAC TTCAACACGC TGATGGCAGT GGTCGGGGGC CTGAGCCACA  
6451 GCTCCATCTC CCGCCTCAAG GAGACCCACA GCCACGTTAG CCTGAGACC  
6501 ATCAAGGTGC CTGGGACTGG GGAGGGGCGG GTGCTTCCCA GGTCTGTCTT  
6551 CACTGGGTCC TCCCAGCAGC ACTGGGGGCT GGGCACAGCT GTCCTCATTT  
6601 GATAGATATG GAAATGGAGG CTCAGAGGGG TTAAGTGCTT TTCTCAGTTT  
6651 GCACAATGGC AACAGCAGAG TGGGGGCTCA CAGGTGCTCA GGGACCCCAA  
6701 AGCTAGTACT TTTTTTTTTT TTTTAAAGAC AGGGTCTCTC TCTCTGTTGT  
6751 CCAGACTGGA GTTCAGTGGT GCAGTCACAA GCTCACTGCA GCCTTGAATT  
6801 CCTGAGCTCA ATCGATCCTC CCACCTCAGC CTCCTGAGTA GCTGGGACTA  
6851 CAGGTGTACG CCACCATGCC TAATTTTGTG ATTGTTATTA ATTTTTTTTT  
6901 TTTTTTTTTT GAGATGGGGT TTTGCCATGT TGCCAGACT GGTCTTGAAC  
6951 TCCTGGGCTC AAGTGATCCG CCTGCCTTGG CCTCCCAAAG TGCTGAGATT  
7001 ATGGCTTGAG CCATTGTGCC TTGCCACTTG TAGTTTCTTC TTTTCTTTCT  
7051 CCTTCATTTT TTATTATTTT TGAAGTATTT TGAAGTATTG AGTAACATAC  
7101 ATATAGAAAA GTATATAAAA ACATATGAGA CTGGGCGTAG TAGCTCACAC  
7151 CTGTAATCCC AGCACTTTGG GAGGCTGAGG TGGGCAGATC ACGTGACATC  
7201 AGGAGTTTGA GACCAGCCTG GCCAACAAGG TGGAAACCCA TCTCTACTAA  
7251 AATACAAAAA TTAGCCAGGC ATGGTGGCAC GCACCTGGAA TCCAAGCTAC  
7301 TTGGGAGGCT GAGGCAGGAG GAGAATTACT TGAACTCAGG AGGCGGAGGT  
7351 TGCAGTGAGC CAAGATTGTG CCACTTCACT CCAGCCTGGG CGACAGAGTG  
7401 AGACTCCATC TAAAAAATAA GAAAAGTATA TAAAAACATA TGAATAGTTT  
7451 AAAGAAAAAT TGTAAAGAAA AACTGTGTA ACTACTGCCC GGGTTGGGAA  
7501 ATAGAACCTT GCCAGGCCCC CAAGCGCCCA GCACTTTAGA GCATAACTCC  
7551 CTCCCCACGA CTTTTGCAAT GATGATCTTG CTTTTCTTTA TAGCTTCACC  
7601 ATGTAGGTAT GCGGTCCAAA ACAATGTGGG GCTTTTGTG GTCTGTTTTG  
7651 AACTTTCTAT GAATGGAATG TTGTTTGTGT TATTTTATGT CTTGCTTTTT  
7701 TCATTCCACA TGGTCTGAG AGTCTTTTCA TTCTGTCATG TGGAGCAATT  
7751 GTTTTTTTCAT TTTTATTGCC ATATAATATT TTATGTACG TCTACCCCAA  
7801 TTCATTTATT TATTTATTTT TTTGAGATGG AGTCTGTCTC TGTCATCCAG  
7851 GCTGGAGTGC GGTGGCGAGA TCTCATCACT GCAACTCCG TCTCCTGGGT  
7901 TTACGTGATT CTCGTGCCCT AGCCTCCTGA GTAGCTGGGA TTATGGGCTC  
7951 GTACCACCAC GTCTGGCTAA TTTTTTGTAG AGACAGGCTT TCACCATGTT  
8001 GCCGAGGCTG GTCTTGAACCT CCTGAGCTCA GGCAATCCAC CCGCTTTAGC  
8051 CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACTGCCCC CAGCCTACCC  
8101 CAATTTTATG ATTGATTCTA TTGTTGAATG TTGGGGTTTT TCCTTTTCTT  
8151 TTCTTTCTTT CTTTTTCTTT CTTTTTTTCT TTTTTTTGGA GAGGGAGTCT  
8201 TGCTCTGTGC CCAGGCTGGA GTGCAGTGAC GCTAATTTGG CTCACTGCAT  
8251 CACTGCACCC TCTGCCTCCC GGGTTCAAGC GATTCTCCTG CCTCAGCCTC  
8301 CTGAGTAGCT GGGACTACAG GCATGCACCA CCACACCCGG CTAATTTTGG  
8351 TATTTTTTTT GTAGAGATGA GGTTTCCACC ATGTTGGCCA AGATGGTCTC  
8401 CATCTCTTGA CCTCATGATC CATCTGCCAT GGCTCCCAA AGTGCTGAGA  
8451 TTACAAGTGT GAGCCACCAC GCCCAGCTGG TTTTTCAGT TTTTGCTGTT  
8501 TGGACGGGGT GGCTGAGTAT GTTCTTCCAG GTCATTGTCC TGTGCTGCCT  
8551 TGCTCCCTG AGCCTCTGTT TCTCCTGTTA AATGTTGATG ATTCCCTGCA  
8601 TCCAGGCCTG GTTTAGAGGT GTGGTGCTTT TGGCAGTGAG TATTGCCTTG  
8651 AATTATGCTG AATGAATTCA ATCCCCAGGG GCTGAGAGAG CCAGTCGTGG  
8701 GGGACAGTAA GGGAGGTTTT TACTCTTTCA CCTGTCCCTG ACCCTGACTC  
8751 CTCCTCACCC CCTCCTACAT TTCCAGGGCT GAGGTAGGGA GGATAGTTGT  
8801 GGGGGTATGA CTCCTCTGTC CTTTGTCCCC AGCTCTGGGA GGGTCTCACG  
8851 GAACTAGTGA CGGCGACAGG CAACTATGGC AACTACCGGC GTCGGCTGGC  
8901 AGCCTGTGTG GGCTTCCGCT TCCCGATCCT GGGTGTGCAC CTCAAGGACC  
8951 TGGTGGCCCT GCAGCTGGCA CTGCCTGACT GGCTGGACCC AGCCCGGACC  
9001 CCGCTCAACG GGGCCAAGAT GAAGCAGCTC TTTAGCATCC TGGAGGAGCT  
9051 GGCCATGGTG ACCAGCCTGC GGCCACCAGT ACAGGCCAAC CCCGACCTGC  
9101 TGAGCCTGCT CACGGTGAGG AGCAGGGGGC AGGGAGGTGG GGAGCTGGGC  
9151 ACCAGGGGTT GACAGTTTCC CCAGGTCCTG GCTGTGGGCG TGGCCTGGGG  
9201 CTCTGGGTTT TGGCCAAGAA ACTGAGATCT AGCGTGGGCT CTGGGGTTTG  
9251 GAGTGGATGC TGAGAAGGGG TCCAGGCTCT GGTGGGGCT GTGGACTGAG  
9301 GTCTGATCTC CAGGCTGGTA TGTGGACTGT GGGCAGTTTG AACTGGGCCT  
9351 GGGTCCCGGG TTGAGTTCTG GCAATGGGCT GTGTTCTAGG GCTGGGCCAA  
9401 GCTCTGCATT CTGTGGGCAG GGGTGGTTTC TAAGCATGGC CCTGGGCTCG

FIGURE 3, page 3 of 12

9451 GAGTGAAGTT CTGGGCTTGG CTTTACACTT GGTCTTGGGG TCTAGGGTGG  
 9501 GAGTTGGGTT CTGGTTTAGA TCCAGACAAG GTTCTAGACA TTGGGCTGGG  
 9551 GCTTAAGTGT TAAGGTTTGG AGTGGATTCT TAGCTGCTTC TGGGCTCTGG  
 9601 AGGGGATCAG GGTTGAAATC AGAGCTTCTG GCTGGGTTCC GACCTGGCTT  
 9651 CTTCCCTGAC ATCTTGGCAA TATGTTGTGT TCAAGGTTTG GGGCCATGCT  
 9701 GTGGTTTGAT CTGTGCGCTG GGATGACATG GGGGTTGCTG TGCTGTGTTC  
 9751 TAAGCCAGGC TTTGTCTCGA GTCTAGCTTC TGACCCGAGC TCTGGCTGAG  
 9801 CTGTGGCCTC TAGGTCGACC TTTGGCCCTG GGCTCTGTGG CCGTGGGCAG  
 9851 GGGCCAGTGG GGGTGATCAG ATCTGTGTGT CCCAGGTGTC TCTGGATCAG  
 9901 TATCAGACGC AGGATGAGCT GTACCAGCTG TCCCTGCAGC GGGAGCCGCG  
 9951 CTCCAAGTCC TCGGTGAGGG GGTACTCCCT CCTCTCCACT CTGCCCTTCC  
 10001 CTCCTGAGAA TCCCAGGATG TGAGGATGGG AAGAGCTCTT AGCAGCCACC  
 10051 TCACCCATCC ATCTTGTAAG ACAGAGGCAT CCTGGGGGTA GGGCAGTAGT  
 10101 GTTGGGCAGA CTTCCCTCTC CCAGGGATTC CCCTCTCTGT TCCCCGGGGC  
 10151 TCTGGGCTCC CCCTGCCTCT GGCCCTAGCT CAGGCCCGAC CATTTCATA  
 10201 GCCAACCCAG CCCACGAGTT GCACCCACC ACCCCGGCCC CCGTACTGG  
 10251 AGGAGTGGAC CTCGGCTGCC AAACCCAGC TGGATCAGGC CCTCGTGGTG  
 10301 GAGCACATCG AGAAGATGGT GGAGGTGAGC TCCTGCGGAG CCTGAGCAGT  
 10351 GTGTGGGGAG AGGCCAGTTT GCCGGAGCAC TGCCCTGGAA GCCAGCACGA  
 10401 GTGTCTCTGT CAAGACCCAG CACTCAGCCC CTAGGAGTCA CAGGGCCTGG  
 10451 CAGGCCAGCT GCACGGGGCT GAAGTGCCCC TGGGTAGGGT GGGGGTGGAG  
 10501 GTATGGAACG GGGGTGGTGT CAGAGACCTC TCTGAGACAC ACCTCATCAA  
 10551 ATGGATGGG AACGTGGGAA GGGACAGGAC CTGATGTCCC CTTTACTCTC  
 10601 CCCTCTTCTG GCTCTGCGTG TCCCTCTGCG TGCCCCAGTC TGTGTTCCGG  
 10651 AACTTTGACG TCGATGGGGA TGGCCACATC TCACAGGAAG AATTCCAGAT  
 10701 CATCCGTGGG AACTTCCCTT ACCTCAGCGC CTTTGGGGAC CTCGACCAGA  
 10751 ACCAGTGAGG AGGGCTGGGG ACCTGGGGGA GAGGGAAGGC AACTCAGCCC  
 10801 ACTTCTGCCT GGGCTTCAGT TTCTTGTGTG CAAGATGAGG TCACTGAGCC  
 10851 AGATGATCTT GGCCTGGGAA GCTGCCAGTG TGGGAAAGGG CACTTGCTTT  
 10901 TGTGGGGAGG AGAGGCTGCC AGCTGTGGAG GCGCAGTGGT ATCTCACAAA  
 10951 TTCAGACAGA TGGGGGGCTC CACCTGAGTC TTGCAAAGAC TGTGACCTGG  
 11001 GGACTGTGGC TACAAAAGTG CTGTTTATT TGTGGAGCTC ACAGCTGTCA  
 11051 AGAAGTGTGG GCAACTTGAG CTCCTGGATA GTCTGTTCTA ATGAATAGAT  
 11101 AAGAAAGGTT TGTAATTAGC AGTACCCAGT TGTTTATCAA CAGTTCATAT  
 11151 GCTGACAATT TGGA AAAACA GCTGGTTCTC TGAAGTAGGT TAAACATGCC  
 11201 CCCTGAAGCC AGATTCATGC CCTATTTTGT CTGAGCAGAA AAAACTCCAT  
 11251 TCAAAATTTA AAGTCCATCT CAGGTCGATT TATTTTAA TGTTACCTGT  
 11301 ATTTCAAAAA TCTGTTGTTT TTTATTTCCA CATTACAAAA ATCCACGGTA  
 11351 AAATAAAATC TAGGTGGTAA AATAAAATTA TAGTGAACAA AATGTTTAAA  
 11401 GTAAGAAGTG AGAGGCCAGG TGCGGTGCCT CACGCCTGTA ATCCTAGCAC  
 11451 TTTGGGAGAC TGAGTTGGCA GGATCAATT AGGCCAGGAG TTTGAGCCCA  
 11501 GCCTGGGCAA CAGAGTAAGA CCCTGTCTCT ACAAAAATTA TTATTATTAT  
 11551 TTTTGAGACA GAGTCTCACT CTGTTGCCCA GGCTGGAGTG CAGTGGTACA  
 11601 ATCTCGGCTC GCTGCAACCT CCACTTCCTG GGTCAAGTG ATTCTCCTGC  
 11651 TTCAGCTTCC TGAGTAGCTG GGATTACAGG CATGCATCAC CGTGCCTGGC  
 11701 TAATTTTTGT ATTTTATGCA GAGATGGGGT TTTACCATGT TGGCCAGGCT  
 11751 GGTCTCAAAC TCTTGACCTC AAGTGATCTA CCTGCCTTGG CCCCCAAAG  
 11801 TGCTAGGATT ACAGGCATGA GCTACTGCTC CTAGCCTAAA AAAATTTTTT  
 11851 TTGGGCATGG GTGGCACGTG CCTGTAGTCC CAGCTACTCA GGAGGCTGAG  
 11901 GCAGGAGGAA CCCTTGAGCC CAGGAGGTTG AGACTGCAGT GAGCTGTCAT  
 11951 CACACCACCT CACTTCAGCC TGGGTGACTG CGCGAGATCA CCCCATCAA  
 12001 AAAAAAAAAA AAAAGAAAAA AAAAGGAAGA AATGAAAGTC CCCTCTTTCC  
 12051 TTTTCCACTG GTAGAAGTTG CCATGATTAA GCACTGTAA CAATATTAAG  
 12101 CTTGGCAGTA TGTGGATTCT TCCAGTCTTC TTTTCCAGG CAGGTGCACA  
 12151 TTGATAGAGA TTTTGTGTTT TTGGTGTCTG TTTTATGGAC AAACAGGATT  
 12201 AGAGCATAAA TCTAGTCTCT CTTGTGGCTT TTATCATAGC TGCTTTATTT  
 12251 CTTCTCCAG ATTTTAGGCA GAGGTAGTTG AGTTCCATGT TTTCTCCCTG  
 12301 GGTGTTGGG TGGATTTTAA TCTAGACCAC CTTTTCAGTG AGAATGACCC  
 12351 TTTGAGACGA TGGAGGCCTC AGCTTCATGC AGCGGGCTCA GCCTTAACCC  
 12401 TCCACCTCCT GCAGGCCCCA AGCTGTGTGT GTGTGTGTGT GTGTGTGTGT  
 12451 GTGTGTGTGT GTGTGTGTGT GTTGGTAAGG GGAAAGCCCC TGGTTGGGTA  
 12501 TCAAAAACCT AGCACCTGGT TCGGCAGGAG GGAGACCAGC ACCGGCTCCC  
 12551 CAGGACCAGG CCCAGCTCAC CACTTCATTG TAAAGCTCCC TCTTTGTTTC

FIGURE 3, page 4 of 12

12601	TGGAAC TTGG	GTGTTTCCAT	TTCTTTCTTA	CAAAATTATC	TATGCATTTA
12651	CAGCAATTGT	TGATATATCT	TTAGGCAGCA	TCTAGGTA CT	TGTA GTGGGT
12701	TCTCTTTTTT	CTTTTTTCTT	TTTTTTAA TC	ACCC TCTCTT	TTTTTTGAGA
12751	CAGAGTCTCA	CTCTGTGCTC	CAGGCTGGAG	TGCAATAGCG	CGATCTTGGC
12801	TCACTGCAAC	CTCTGCC TCC	CAGGTTCAAG	TAATTCTCAT	GCCTCAGCCT
12851	CCCAAGTAGC	TGAGATTACA	GGCACTGGCC	ACCAGACCCG	GCTAATTTTT
12901	TTTTCTTTTT	CTTTTTTTTG	AGACGGAGTT	TCGCTCTTTG	TTGCCCAGGC
12951	TGGAGTACAG	TGGTGTGATC	TCGGCTCACT	GCAACCTCCG	CCTCCCGGGT
13001	TCAAGTGATT	CTCCTGTCTC	AGCCTCCCGA	GTAGCTGGGA	TTACAGGCGC
13051	GCGCCACCAT	GCCTGGCTAA	TTTTGTATTT	TTTTTTTTTT	GAGACAGAGT
13101	CTCACTCTGT	CACCCAGACT	GGAGTGCGGT	GGCGCGATCT	CGGCTCACTG
13151	CAAGCTCTGC	TTCCCGGGTT	CATGCCATTC	TCCTGCCTCA	GCCTCCGGAG
13201	TAGCTGGGAC	TACAAGCACC	CACCACCGTG	CCCGGCTAAT	TTTTTG TATT
13251	TTTAGTAGAG	ACGGGGTTTC	ACCGTGGTCT	CGACCTCCAG	ACCTCGTGAT
13301	CCACTAGCCT	CAGCCTCCCA	AAGTGCTGGG	ATTACAGGCG	TGAGCCACCT
13351	CACCCAGCCT	AATTTTGTAT	TTTTAGTAGA	GATGGGGTTT	CACCATGTTG
13401	CGCAGGCTGG	TATTGAACTT	CTGACCTCAG	GTGATCCGCC	CGCCTCGGCC
13451	TCCCGAAGTT	CTGGGATTAT	AGGCGTGAGC	CACCGCACCT	GGCCTAATTT
13501	TTGTATTTTT	AGTAGAGATG	GAGTTTTACC	TTGTTGGCCA	GGCTGGTCTT
13551	GAACTCCTGA	CCTCACCTCA	GGTGATCTGC	CCACCTCGGC	CTCCCAAAGT
13601	GCTGGGATTA	CAGGCATGAG	CCACTGTGCA	CCCGGCCTAA	AAATCACCAT
13651	CTTGACAGAA	CTTCACGCCT	TGCTTTTTGT	TTTTTTTCAT	CTTTGTGCTT
13701	GTTTTCCACT	TAACCC TTGA	TCACAGACAT	CTTTCCATGT	GGATTCATGT
13751	AGAACTACCT	CATTCTTTAG	AACAGCTGCA	GAGTATTCCA	CTGTGCGGTT
13801	AGTCCATCAT	TTCCCTAACC	ATCCTCCTGC	TGATGGACAG	TTAGACTGTT
13851	CCAGTTTTTC	AGTATGATTC	TATGCCAGGC	TGCCATGAAC	GTCTTTTAC
13901	TGATCCACTC	AGGCCAGTAT	TTCTGTAGGA	GAAATTCCTA	GAAGTGGGAT
13951	AATTGGATCA	AAAGATATGC	ACATTCTAAA	TTAGGAGAGA	GACTGCCAAA
14001	CTGACCTCAG	ACAAGGTTGT	ACCAGTTTGC	ACCCCCATCA	GCAGCGTACA
14051	AGTGCCTGGT	TCCCAACTTC	CTCGCCAACA	GGGATGCTAT	AAAAAGCTTC
14101	ACAATTTTGC	CAGTCTCATT	GGCAAATGGT	ATCTTGGTTA	AATTTGCATT
14151	TCTTTAATAC	TAAGTGGGGG	TAGGGTATCT	TTTCATATGT	TTATTGGCCA
14201	TTTATTTCTT	CTGTCAATTG	CCTGTTCTGA	TTCCTTGTCC	ATTAT TCTAC
14251	TGGGTTTGT	GGTCTTTTTC	TCATTGATTT	TTAGAACTC	TGTTAATGGA
14301	TATTAACCCT	TTGCTGTTGA	ATGTGTTTGC	AAATATTTTC	TCCCTGTCTG
14351	TCATTTATGT	GTCTTTTTCC	ATATAAATTT	AAAAAATTTT	GGTGGGCTCA
14401	ATAGGTCAGT	CTTTCCCTTC	CGGGCTTCTG	GGATTTGTGT	TCGGGGTAGA
14451	AAGGCCCTCA	GCCCTCAAG	ATTATAAAAT	TATAAAACCT	TTTCTTTTTT
14501	TTTTTTTTTT	CTGAGACAGG	GTGTCTTGCC	ATGTCACCCA	GGCTGGAGTG
14551	CAGTGGCATG	ATCTTGGCTC	GCTGCAACCT	CCACCTCCCA	GGTTC AAGTG
14601	ATTCTCGTGC	CTTAGCCTCC	CGAGTAGCTG	GGATTATAGG	TGCCCTGCCAC
14651	TATGCCTGGC	TAATTTTTTG	TATTTTTAGT	AGAGACGGGG	CTTTGCCATG
14701	TTGGCCAGGC	TGGTCTTGAA	CTCCTGACCT	CGTGATCCAC	CCGCCTTGGC
14751	CTCCCAAAGT	GCTGGGACTA	CAGGCGTAAG	CCACTGTGCT	CGGCCCTATA
14801	TTTTTTTCAG	ATAGCCAGTT	ATCCTAATGC	TCCCTTGATT	TGATGGACCA
14851	CCTGGATCAC	ACATTATGAG	CCCCCTCATA	AGCAGGTGGG	AGTCTCAAGC
14901	GAGGGCCAGT	CCCGATGGGA	ATAGCACTTG	GTGGCTGAGG	ACCCTCCTAT
14951	CTGTGCAGAC	ACTGTTGTAA	AACTTCACAT	GCATCATCTA	ATTTAGTCCT
15001	CACCAAAATC	CTATGAAATG	TAGGAATGAT	CATTACACCC	ATTTATAGAT
15051	AAGGAAACGG	AGGGACAGGG	AGATTACTCC	GCTACAGGTC	AAGAGGCAGG
15101	GAAGTAGAGC	TGCGATTTGA	ACTGAGGTCT	GTGTCTAGAA	CACGTGCTCA
15151	TTCTTTCCCT	AAAATGTATT	CATAGGTGAA	AAAGGGCTTC	TGCGGAAAGC
15201	CCTGGGTTAT	GTGGGAAACC	CTGGATTTAC	AGCTGTCTTT	CCAGCAGGAT
15251	GATGCAGGAG	AGAGAGGGAT	GCGATTTCTC	CCAATCTCTC	CTGGTCCCAG
15301	AACTCATTAG	AGAGTTCTCC	CTGCTGAGGG	CTCCCGACTG	GTGTTGCACA
15351	CAGTACACTT	CGGGAGCCCG	AGGCTGATGG	TTCCATGGAA	AGTACACAGT
15401	CATTTTAGTT	TGCACACCAA	GTGTGAAGTG	GGCAGGACAG	GCCACTGTTC
15451	TGAGAAGGAA	CCCAGGGAAA	GGGACTGGCC	CAAGACCACA	CACTGGTTAG
15501	CGGCACTTCC	CACATCTGCC	TGACCCCTAG	TCCAGTGCCG	CCTTTTCTTT
15551	ACTCTGCAAC	AGGAGTCCAA	AATCAGGAGT	TCCATGAGGA	CACTGGGAAC
15601	AGTGGGATGG	GTTAGGCCAG	CGGTGGATGG	TTCTGGGGAG	GGCCCGAGCT
15651	GAAGCGCCCC	CGCAACTCCC	CACAGGGATG	GCTGCATCAG	CAGGGAGGAG
15701	ATGGTTTCCT	ATTTCTCTCG	CTCCAGCTCT	GTGTTGGGGG	GGCGCATGGG

FIGURE 3, page 5 of 12

15751 CTTCTGTACAC AACTTCCAGG AGAGCAACTC CTTGCGCCCC GTCGCTGCC  
15801 GCCACTGCAA AGCCCTGGTG AGAGTCCCTT TCCCGGCTCA CGGCCCAAGC  
15851 CACGCCCTTC CAGCCCCGGC CCGGCCCTCC CTTCTGGCCC CGCCTCTGCC  
15901 AGAGCCCTTC TCAAGCCAGG AAAACCTGGT AATTCTATTT GCCTCTCCTC  
15951 CTGTGGTTCT GCGCGGGGCC CTGAGGCGGG CTCTAAAGCC CTAGTCTCAC  
16001 CCTCAAGAAG GAAGAAGTAG AGTCATCACC TCTAAATCCC TCCTCCCACC  
16051 ACGGCCCTC CTCTATTGCA GATCCTGGGC ATCTACAAGC AGGGCCTCAA  
16101 ATGCCGAGGT GAGATGGAAT GACTGGAAGG GCTGCTGGGC AGTGTTTTTT  
16151 TTGTTTGTGT GTTTGTTTGG GAGAGTTACT ATTTTGGTGG GGCAATTGCC  
16201 AAGGAGTGAA GTACCTTAAA ATCAGAGGCG CATGGCCGGG CATGGTGGCT  
16251 CAAGCCTGTA ATCCCAGCAC TTTGGGAGGC CGAGGCGCGC AGATCACCTG  
16301 AGGTCAGGAG TTCAAGACCA GCCTGACCAA CATAGCGCAA CCGCGCTCT  
16351 ACTAAAAATA CAAAAAGTAG CTGGGCGTGG TGGCACCCAC CTGTAATCCC  
16401 AGCTACTTGG GAGGCTGAGG CATGAGAATC GCTTGAACCT GGGAGGCGGG  
16451 GTTTGCAGTG AGCCGAGATC ACGCCACTGC ACTCCAGCCT GGGCAACAGA  
16501 GAGGGCTCTG TCTCAAAAAA AAAAAACAAC AAAAAACCC CAAAAACCAA  
16551 AACCCACAAA AATCAGAGGC TCAAGATGAC TGATGTGAAG GGAGTGGCGT  
16601 TTAAGAGGCC ATTTATTTTG ATGACGCAGC TGCCCAGGAA CAGAGAACAT  
16651 GGGAGAAGGC ATAGACTGAC AATTAGGAGG AGGAGAACAC TTTGGAAGGA  
16701 GACTCTTATT TTGGTGGGGC AGCTGCTCAG GAACAAAGGT TCCTGGTAGG  
16751 GGGGCGCAAG CCTGCGGGAT GGGATGGAGG GTATTCTGAC CAATGTCCCT  
16801 GGCTGGCTCT CCATTGTCTC TCCCCAGCC TGTGGAGTGA ACTGCCACAA  
16851 GCAGTGC AAG GATCGCTGT CAGTTGAGTG TCGGCGCAGG GCCCAGAGTG  
16901 TGAGCCTGGA GGGGTCTGCA CCCTCACCTT CACCCATGCA CAGCCACCAT  
16951 CACCGCGCCT TCAGCTTCTC TCTGCCCCGC CCTGGCAGGC GAGGCTCCAG  
17001 GCCTCCAGGT AAGAGGGAGT CATTCTGTAC TGGCCTGTGG AGGGAAGGAT  
17051 GCAGGGCTAC TGGGGCAAAG AACGCAGGAT GGAAGCCATT CCAAAGTGCA  
17101 TAATTCTCTT TTTGTGGTGG GATAATAAAG AAGGGACAGG CCGGGCGCGG  
17151 TGGCTCACGC CTGTAATCCC AGCACTTGG GAGGCGGAGG CGGGCGGATC  
17201 ACGAGGTCAG GAGATCGAGA CCATCCTGGC TAACACGGTG AAACCCCATC  
17251 TTTACTAAAA ATACAAAAAA AAAAAATTAG CCAGGCGTGG TGGCGGACGC  
17301 CTGTAGTCCC AGCTACTTGG GAGGCTGAGG CAGGAGAATG GCATGAACCC  
17351 GGGAGGCGGG GCTTGCAGTG AGCCGAGATC GCGCCACTGC ACTCCAGCCT  
17401 GGGCGATAGA GCAAGACTCC GTCTCAGAAA AAAAAAAAT AAAAAATAAA  
17451 GAAGGGACAG GTAAGGGTGC CAGAAAGTGG CCAGGAAGCC CTGGACCTTC  
17501 TGAGGCTGAG GAGAGAGACC CTAATTTATA AAGAGGTATA AAAGTGAAAG  
17551 AGGCTTCAAG ATTCCAGTTA CAGTCTTATT TTGTTGGAGG GGTAAACAAA  
17601 GGATTGGAGA AGGTGTTATA TGAGCCATTG GCTTGCCTTT CCCTTCTCTG  
17651 CTGCTCTGGA GGCTCTTCTG GGGAAAGTCC CTTGCCCTGA TAATGTCTCTG  
17701 GCAGCTCTCT TGGGGTATTT GATGGTTTGA GGTCAGTTTG CTGAATGACA  
17751 ACTGGCCAAA TGATTATTTT GCTGAGAACG GTCCGAACAA CTATGTTAAA  
17801 CTGGGGTCTA AGGTAGTTGA TCACAACGTG TTGGGTGGC ATAAGTCCTC  
17851 AAAAAACAGA GGCAGGCACA GGCATACAT CCTCAAAAAT AGAAAAGATA  
17901 AATCCATTTG CATTGAGCCT TCCAGAAGTG CTGGGGTCTA AAATGTGAAA  
17951 TACACACAAA ATTGACATTT AAGCAAATG CGCTGACAAA TCTGTGGCTG  
18001 AAAAAGCTGT GGCAAAACAA AAACATAGAA AAAGAGCCTC AAAAATTGGG  
18051 CTGAGGCCAG GCATGGTGGC TCACGCCTGT AATCCTAGCA CTTTGGGAAG  
18101 CCAAGGTGGG TGGATCACCC GAGGTCAGGA GTTGAGAGCC AGACTGGCCA  
18151 ACGTGGCAAA ACATCATCTC TACAATACAA AAATACAAA ATTAGCTGGG  
18201 CGTGGTGGCA GGCGCCTGTA ATCCCAGCTA CTTGGGAGGC TGAGGCACGA  
18251 GAATCGCTTG AACCTGGGAG GTGGAGGTTG CAGAGAGCCG AGATTGCGCC  
18301 ATTGCACTCC AGCCTGGGCG ACAGAGAGAG ACTCTGTCTC AAAAAAAAAA  
18351 AAAAAAAAAA AAATTGGGCT GTGAGGTCAT GCAGGGAATT GATTTTTGGT  
18401 GGGTGGGTCT GCTTCTGGGA TGATGTGGAT GCCTCCCGTG GAGAGGGGAA  
18451 GGGTTGATGA AGTCCAGGG AGCTGGAAGT GTGTTCTGCA GCAATCCCCC  
18501 TCCAGCAGA GATCCGTGAG GAGGAGGTAC AGACGGTGGG GGATGGGGTG  
18551 TTTGACATCC ACTTGTAATA GATGGTGAGT CCTCCACAG CTGGCACCAG  
18601 AGCTCCCCAC TGAGGGCTGG GGGGGAGCTG GGGAGTATCA GGGAAATGGG  
18651 TGCTTTATCC AAATGGCTCC AAGCCAGGTG GGCTACTACC TTGTTGTTAG  
18701 GGGGGTGTCT TCCTCACAACT GTGTTTTTCT CTTCCAGCT GTGGTTGGAT  
18751 CAAGGACTCA TTCCTGCCTT GGAGAAAAATA CTTCAACCA AGCAGGGAGC  
18801 CTGGGGGTGT CGGGGCAGGA GGCTGGGGAT GGGGGTGGGA TATGAGGGTG  
18851 GCATGCAGCT GAGGGCAGGG CCAGGGCTGG TGTCCCTAAG GTTGTACAGA

FIGURE 3, page 6 of 12

18901 CTCTTGTGAA TATTTGTATT TTCCAGATGG AATAAAAAGG CCCGTGTAAT  
 18951 TAACCTTCAC CATCAGCGCC TAGAATCCCG GGGGGTAGGG GGATGGTATA  
 19001 CTTTACAGGA TGACAATCTT GGGAGCTAGA ACTTTGTAGC CAGAGAACT  
 19051 TGGGAGGTCT GGAATCTCAT GTGTCTGGAG TCTTGGGGAA GAGAATCTTA  
 19101 GAAGCAGAAA ACCTTGGAAC ATAAGAATCT TGGGGAGGGT CTAGGATCTT  
 19151 GAGGAGACCA GATCCTTGGA CATCTAAAAC TTGAAACTAG TAGGTCTGCA  
 19201 CCCGAGAAAT GCAGGGCCAG TCATGCATAC CCAAAGCCTT CAGCCCATGG  
 19251 CCGAAATTCC CTTGCTGGAC AGGGGGCCTT TCAGCCCCTG CTGGGACGCT  
 19301 TCCAGTAACA GGGCCCTCAC TGCAGGAATC GTGGGAGGGA GAGGGGCAGC  
 19351 ACAGAGTTGC TGGCTGTCGG GGAAGGGAGG GAGGGCCCTG GGCAGTCCGA  
 19401 GGGCCCTGCT GGGCTTGTGC CTCAGGGTGG GGGCTGCACT CCTCCGCCTT  
 19451 GCAGCCTCCT GGCCTGGTGC TGCTGCCAGC CGGAAGGACA GTGACTTCCA  
 19501 GAGGAAATGC ATATTGATCC TGCTTTCAGC CTCCGGTGGT GGCTTCTCCC  
 19551 AACCAGCTC TTCCCTCCTG AGCCTGCAGC ACGGAGGTTT TGGGGGTAC  
 19601 TGCTACCTAA AGAAGGCTAA GGCCACTTCT GAGGCTGGTC TGGGAGTTTA  
 19651 CTAAAGGTTT TGAAGCTGGG CCGGGCTGCC CCTGGGATCA GGAGACTCCA  
 19701 GACAGCAGTC CTGACAATGG GAACTACCTC CTCAGTCCCC CAACTGGGA  
 19751 GGTGTCCAC AGCAGCTGTA GGATTGTCT AGGGGTGGAG ACCTGAGCAC  
 19801 CTTCCACTCC AAAGCACAGT ATCTGTGGGC CTGGCAGTGG CCTCAGTTCC  
 19851 CCCATGAGTG CCCCGTCCC CCACCCAGG GTTTCCTTCC ATCAGTCCA  
 19901 TCCCTGCTTT GAGACCCAC TCCCCCTGGC CTGTCTTTTA TTTTGGGTCA  
 19951 CTCCCTTCTC TTTCTGGTC ATATCTCTCC TGCAGGCCTA CCCTGTGTTG  
 20001 GGGCCCCCAG CCTGTCTCT GCATCGGGTG CCCCCCTGCC CCTCCTTCTG  
 20051 TCCTCAGCCC CCTCCGCCCT TCCCCCTCTT GAGGCTGTAA TATCCGTTTC  
 20101 ACGATTTGGG GGCTGAGTTG CTATAACAAC AGACGGCGAT TGTGTTGTGA  
 20151 AGAGCAGCTC GCTCCTGTGC CGCCTGCCTC CTGTGCTGCC TCCATCCCTG  
 20201 CAGCCAGTCC GGTTCCTCTT GGCTCCTCTC GTCATACCC TCCAGTTCCA  
 20251 GTCTGGCCTC TTCCTGGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTAT  
 20301 GCATGCATGC ATATGTGTGT CCAGGTCTGC CTGCCCCGGA TGTGACAAGT  
 20351 AGCGGTCTTC ATGGTTGCAT GTGTCTGAAT TTGGTGTCTG AGCTTCACAT  
 20401 TGTATGCGCC TGTGTGCATG TGTGTGCATG GACATGCATG CTGTATCTGC  
 20451 TGTGTTTCCC CTCCCCCATG TGTCCCCACT GGCTTTTGA CATGGGAGAA  
 20501 GGGCATGTGC TCAGCATATC ACTCAACTGT CCACATGGG TGGGTACCTG  
 20551 TGTGTGGTGT GTGTGTGTGG GGGGTGTGTC TTGAAGTGGC AGGTCCCAAA  
 20601 TGCTTAGGCA ATCTGAACCT TGGACCTTGC AGAGAGGAGA GATGTCCCTG  
 20651 TAGGTGGGAG GGACAGGGAG ATGCAGCAGC TGCCCGGTGA CCTTTCTGC  
 20701 CTTGATGGG CAAAGCTGGG GGTAGGGAAA GGAGACAAGT GCTCATACTT  
 20751 ACCTCCCTCC CTGCCAGGC TCCTCTGTAA GGGTCTGAGT CTGTCTCTGT  
 20801 GAGCCATTGC ATCTGTCTGT CTATGCCCTG ATGCCTGGAT GGACAAGGGG  
 20851 TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG AGTGTGAGGC TGCAGGAAGA  
 20901 GGAACAGTGG GGGATGGGCA GGAAAGTGGG CTGTGGGGTC AGGGAGGCGA  
 20951 T (SEQ ID NO:3)

# **FEATURES:**

Start: 3000  
 Exon: 3000-3072  
 Intron: 3073-3753  
 Exon: 3754-3855  
 Intron: 3856-4363  
 Exon: 4364-4427  
 Intron: 4428-4786  
 Exon: 4787-4918  
 Intron: 4919-5702  
 Exon: 5703-5853  
 Intron: 5854-6056  
 Exon: 6057-6230  
 Intron: 6231-6389  
 Exon: 6390-6506  
 Intron: 6507-8832  
 Exon: 8833-9114  
 Intron: 9115-9885  
 Exon: 9886-9963



Intron: 9964-10201  
 Exon: 10202-10324  
 Intron: 10325-10638  
 Exon: 10639-10754  
 Intron: 10755-15675  
 Exon: 15676-15817  
 Intron: 15818-16071  
 Exon: 16072-16108  
 Intron: 16109-16828  
 Exon: 16829-17008  
 Intron: 17009-18491  
 Exon: 18492-18565  
 Stop: 18566

**CHROMOSOME MAP POSITION:**  
 Chromosome 11

**ALLELIC VARIANTS (SNPs):**

DNA			
Position	Major	Minor	Domain
5539	C	G	Intron
5658	T	G	Intron
5861	C	T	Intron
6023	A	G	Intron
6799	C	T	Intron
9579	C	A	Intron
9842	T	C	Intron
10159	T	C	Intron
12025	A	- G	Intron
14723	T	C	Intron
14996	G	A	Intron
16153	T	G	Intron
16181	G	A	Intron
16756	A	G	Intron
18059	A	G	Intron
18364	A	- T	Intron
18861	G	A	Beyond ORF(3')
20443	G	A	Beyond ORF(3')
20881	A	T	Beyond ORF(3')

Context:

DNA

Position

5539 AGACTCAGTTCAGGCATGAAGTCTCCGTGGGCTCTGAGGGTTCGGGGCTCTTCCGGGGTA  
 GAATTTGTCGTTCCACCTCTGTTTTCCATGGCACTTTGTACAGACTCCTGTACAAAGAC  
 CTCTGTACATGTGTACGCTGTTTTGTGATCATGTGTTTCTGTGTCTGTCTCCCTCAGTA  
 GACTGTGAGCTCCTCGAGGGCAGGAACCGTGTCTTACTCATCTCTGTATTCCCAGCGCCT  
 AGCACAGTGCCTGGCACAGAGTACGTTGTTTATAAATGTGTGTTGAGTGCATGACGGGGT  
 [C, G]  
 GGGGGAGATGAGGAGGAGTTGCTGGGACTGGGAACATTTCGTGCCTAGGACAGTGCCTCGC  
 ATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTCTGTGTGAGTGGGGGGCCACGAGGC  
 ATGCGCATGTCCAGCAAAGGGCTCACTACCCCTGCCCCCAGCCCTACCTACAAGTGGGA  
 AGCGGCAGGTGACTCAGCGGAACCGTGTGGGACAGAAAAGCGCAAGATGTCCCTGTTGT  
 TTGACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCT

5658 CCTCTGTACATGTGTACGCTGTTTTGTGATCATGTGTTTCTGTGTCTGTCTCCCTCAGT  
 AGACTGTGAGCTCCTCGAGGGCAGGAACCGTGTCTTACTCATCTCTGTATTCCCAGCGCC  
 TAGCACAGTGCCTGGCACAGAGTACGTTGTTTATAAATGTGTGTTGAGTGCATGACGGGG  
 TGGGGGAGATGAGGAGGAGTTGCTGGGACTGGGAACATTTCGTGCCTAGGACAGTGCCTC  
 GCATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTCTGTGTGAGTGGGGGGCCACGAG

[T, G]  
CATGCGCATGTCCAGCAAAGGGCTCACTACCCCTGCCCCCAGCCCTACCTACAAGTGG  
AAGCGGCAGGTGACTCAGCGGAACCCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTG  
TTTGACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTACCTACTTGGAGTATCGCTCC  
TTCTGCAAGATCCTGGTGCGGCCGAGGGCTGGGGGTGAGGGTCCAATGTGGGCTGGA  
AGAGAGTTCTAGGAGGGGAGGGTCCCTGGCGTAGGCTGGGTACAGGGTGCATCAGGGG

5861 CTGGGACTGGGAACATTCTGTCCTAGGACAGTGCCTCGCATTATGTAGGTTCTCAGTAAG  
CGTGAATGGTGTCTGTGTAGTGGGGGGCCACGAGGCATGCGCATGTCCAGCAAAGGG  
CTCACTACCCCTGCCCCCAGCCCTACCTACAAGTGAAGCGGCAGGTGACTCAGCGGA  
ACCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGTTTGGACCCTGGAGCCCATGG  
AGCTGGCGGAGCATCTACCTACTTGGAGTATCGCTCCTTCTGCAAGATCCTGGTGCGG  
[C, T]  
CGAGGGCTGGGGGGTCAGGGGTCCAATGTGGGCTGGAAGAGAGTTCTAGGAGGGGCAGGG  
TCCCTGGCGTAGGCTGGGTACAGGGTGCATCAGGGGTTTCAGTGTAACCACTGAAGGTC  
AGCTGGAGGGTGAGGAGTGGCTATCAGTGAGGGGAGAGGCCGCAAGGTGCTGAGGCCAC  
TCCTCATGCCCCAGTTTTCAGGACTATCACAGTTTCGTGACTCATGGCTGCACTGTGGAC  
AACCCCGTCTGGAGCGGTTTCATCTCCCTCTTCAACAGCGTCTCACAGTGGGTGCAGTCT

6023 GGCAGGTGACTCAGCGGAACCCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGTTTG  
ACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTACCTACTTGGAGTATCGCTCCTTCT  
GCAAGATCCTGGTGCGGCCGAGGGCTGGGGGTGAGGGTCCAATGTGGGCTGGAAGAG  
AGTTCTAGGAGGGGAGGGTCCCTGGCGTAGGCTGGGTACAGGGTGCATCAGGGGTTTC  
AGTGTAACCACTGAAGGTGAGCTGGAGGGTGGAGTGGCTATCAGTGAGGGGAGAGGCC  
[A, G]  
GCAAGGTGCTGAGGCCACTCCTCATGCCCCAGTTTCAGGACTATCACAGTTTCGTGACT  
CATGGCTGCACTGTGGACAACCCCGTCTGGAGCGGTTTCATCTCCCTCTTCAACAGCGTC  
TCACAGTGGGTGCAGCTCATGATCCTCAGCAAAACCACAGCCCCGAGCGGGCCCTGGTC  
ATCACACACTTTTGTCCACGTGGCGGAGGTGCCTGCCCTCCCTCCCGGTGTCTCCCAACC  
ACCCACATGCCAGTCAGGCCAACCCCTTCCCTTCCCCTAACCCACTGCCTTCTCTCTAGA

6799 CCATCAAGGTGCCTGGGACTGGGGAGGGGCCGGTGCTTCCCAGGTCTGTCTTCACTGGGT  
CCTCCCAGCAGCACTGGGGGTGGGCACAGCTGCCTCATTTGATAGATATGGAAATGGA  
GGCTCAGAGGGGTTAAGTGTCTTTTCTCAGTTTGCAATGGCAACAGCAGAGTGGGGGT  
CACAGGTGCTCAGGACCCCAAAGCTAGTACTTTTTTTTTTTTTTTAAAGACAGGGTCTC  
TCTCTGTGTGTCAGACTGGAGTTCAGTGGTGCAGTACAAGCTCACTGCAGCCTTGAA  
[C, T]  
TCCTGAGCTCAATCGATCCTCCCACCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGTAC  
GCCACCATGCCTAATTTTGTATTGTTATTAATTTTTTTTTTTTTTTTGTAGATGGGG  
TTTTGCCATGTTGCCAGACTGGTCTTGAACCTCTGGGCTCAAGTGATCCGCCTGCCTTG  
GCCTCCCAAAGTGCTGAGATTATGGCTTGAGCCATTGTGCCTTGCCACTTGTAGTTTCTT  
CTTTCTTTCTCCTTCATTTTTTATTATTTTTGAAGTATTTTGAAGATTGAGTAACATA

9579 CTGGTTGGGGCTGTGGACTGAGGTCTGATCTCCAGGCTGGTATGTGGACTGTGGGCAGTT  
TGAAGTGGGCCTGGGTCCCGGGTTGAGTTCTGGCAATGGGCTGTGTTCTAGGGCTGGGCC  
AAGCTCTGCATTCTGTGGGCAGGGGTGGTTTCTAAGCATGGCCCTGGGCTCGGAGTGAAG  
TTCTGGGCTTGGCTTTACACTTGGTCTTGGGGTCTAGGGTGGGAGTTGGGTTCTGGTTTA  
GATCCAGACAAGGTTCTAGACATTGGGCTGGGGCTAAGTGTTAAGGTTTGGAGTGGATT  
[C, A]  
TTAGCTGCTTCTGGGCTCTGGAGGGGATCAGGGTTGAAATCAGAGCTTCTGGCTGGGTTT  
CGACCTGGCTTCTTCCCTGACATCTTGGCAATATGTTGTGTTCAAGGTTTGGGGCCATGC  
TGTGGTTTGTATCTGTGCGCTGGGATGACATGGGGGTGCTGTGCTGTGTTCTAAGCCAGG  
CTTTGCTCTGAGTCTAGCTTCTGACCCGAGCTCTGGCTGAGCTGTGGCCTCTAGGTCGAC  
CTTTGGCCCTGGGCTCTGTGGCCGTGGGCAGGGGCCAGTGGGGTGATCAGATCTGTGTG

9842 TGGGCTGGGGCTTAAGTGTTAAGGTTTGGAGTGGATTCTTAGCTGCTTCTGGGCTCTGGA  
GGGGATCAGGGTTGAAATCAGAGCTTCTGGCTGGGTTCCGACCTGGCTTCTTCCCTGACA  
TCTTGGCAATATGTTGTGTTCAAGGTTTGGGGCCATGCTGTGGTTTGTATCTGTGCGCTGG  
GATGACATGGGGGTGCTGTGCTGTGTTCTAAGCCAGGCTTTGTCTGAGTCTAGCTTCT  
GACCCGAGCTCTGGCTGAGCTGTGGCCTCTAGGTCGACCTTTGGCCCTGGGCTCTGTGGC  
[T, C]  
GTGGGCAGGGGCCAGTGGGGGTGATCAGATCTGTGTGTCCAGGTGTCTCTGGATCAGTA  
TCAGACGGAGGATGAGCTGTACCAGCTGTCCCTGCAGCGGGAGCCGCGCTCCAAGTCCTC

FIGURE 3, page 9 of 12

GGTGAGGGGGTACTCCCTCCTCTCCACTCTGCCCTTCCCTCCTGAGAATCCCAGGATGTG  
 AGGATGGGAAGAGCTCTTAGCAGCCACCTCACCCATCCATCTTGTAGGACAGAGGCATCC  
 TGGGGGTAGGGCAGTAGTGTGGGCAGACTTCCCTCTCCCAGGGATTCCCCTCTCTGTTC

10159 GGGGGTGATCAGATCTGTGTGTCCCAGGTGTCTCTGGATCAGTATCAGACGGAGGATGAG  
 CTGTACCAGCTGTCCCTGCAGCGGGAGCCGCGCTCCAAGTCTCGGTGAGGGGGTACTCC  
 CTCTCTCCACTCTGCCCTTCCCTCCTGAGAATCCCAGGATGTGAGGATGGGAAGAGCTC  
 TTAGCAGCCACCTCACCCATCCATCTTGTAGGACAGAGGCATCCTGGGGGTAGGGCAGTA  
 GTGTTGGGCAGACTTCCCTCTCCCAGGGATTCCCCTCTCTGTTCCCGGGGCTCTGGGCT  
 [T, C]  
 CCCCCTGCCTCTGGCCCTAGCTCAGGCCCCGACCATTTCATAGCCAACCAGCCCCACGAGT  
 TGCACCCCAACCACCCCGGCCCGGCTACTGGAGGAGTGGACCTCGGCTGCCAAACCCAAG  
 CTGGATCAGGCCCTCGTGGTGGAGCACATCGAGAAGATGGTGGAGGTGAGCTCCTGCGGA  
 GCCTGAGCAGTGTGTGGGGAGAGGCCAGTTTGCCGGAGCACTGCCCTGGAAGCCAGCACG  
 AGTGTCTGTTCAGACCCAGCACTCAGCCCCCTAGGAGTCACAGGGCCTGGCAGGCCAGC

12025 TGGGGTTTTTACCATGTTGGCCAGGCTGGTCTCAAACCTCTTGACCTCAAGTGATCTACCTG  
 CCTTGGCCCCCAAGTGCTAGGATTACAGGCATGAGCTACTGCTCCTAGCCTAAAAAAA  
 TTTTTTTTGGGCATGGGTGGCACGTGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAG  
 GAGGAACCCTTGAGCCCAGGAGGTGAGACTGCAGTGAGCTGTATCACACCCTGCACT  
 TCAGCCTGGGTGACTGCGCGAGATCACCCCCATCAAAAAAAAAAAAAAAAAAGAAAAAAAAA  
 [A, -, G]  
 GAAGAAATGAAAGTCCCCTCTTTCTTTTCCACTGGTAGAAGTTGCCATGATTAAGCACT  
 GTTAAACAATATTAAGCTTGGCAGTATGTGGATTCTTCCAGTCTTCTTTTCCCAGGCAGGT  
 GCACATTGATAGAGATTTTGTGTGTGGTGTCTGTTTCATGGACAAACAGGATTAGAGC  
 ATAAATCTAGTTCTGCTTGTGGCTTTTATCATAGCTGCTTTATTCTTCTCCAGATTTT  
 AGGCAGAGGTAGTTGAGTTCATGTTTTCTCCCTGGGTGGTGGGTGGATTTTATCTAG

14723 GGCTTCTGGGATTTGTGTTGCGGGTAGAAAGGCCCTCAGCCCCCAAGATTATAAAATTA  
 TAAACCTTTTTCTTTTTTTTTTTTTTCTGAGACAGGGTGTCTGCCATGTACCCAGG  
 CTGGAGTGCAGTGGCATGATCTTGGCTCGCTGCAACCTCCACCTCCCAGGTTCAAGTGAT  
 TCTCGTGCCTTAGCCTCCCGAGTAGCTGGGATTATAGGTGCCTGCCACTATGCCTGGCTA  
 ATTTTTGTATTTTAGTAGAGACGGGGCTTTGCCATGTTGGCCAGGCTGGTCTTGAAC  
 [T, C]  
 CTGACCTCGTGATCCACCCGCCTTGGCCCTCCCAAAGTGCTGGGACTACAGGCGTAAGCCA  
 CTGTGCTCGGCCCTATATTTTTTTCAGATAGCCAGTTATCCTAATGCTCCCTTGATTGTA  
 TGGACCACTGGATCACACATTATGAGCCCCCTCATAAGCAGGTGGGAGTCTCAAGCGAG  
 GGCCAGTCCCGATGGGAATAGCACTTGGTGGCTGAGGACCTCCTATCTGTGCAGACACT  
 GTTGTAAAACCTCACATGCATCATCTAATTTAGTCTCACCAAATCCTATGAAATGTAG

14996 CCATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTCGTGATCCACCCGCCTTGGCCCTCCC  
 AAAGTGCTGGGACTACAGGCGTAAGCCACTGTGCTCGGCCCTATATTTTTTTCAGATAGC  
 CAGTTATCCTAATGCTCCCTTGATTTGATGGACCACCTGGATCACACATTATGAGCCCCC  
 TCATAAGCAGGTGGGAGTCTCAAGCGAGGGCCAGTCCCGATGGGAATAGCACTTGGTGGC  
 TGAGGACCCTCCTATCTGTGCAGACACTGTTGTAAAACCTCACATGCATCATCTAATTTA  
 [G, A]  
 TCCTCACCAAATCCTATGAAATGTAGGAATGATCATTACACCCATTTATAGATAAGGAA  
 ACGGAGGGACAGGAGATTACTCCGCTACAGGTCAAGAGGCAGGGAAGTAGAGTCTGCGAT  
 TTGAACTGAGGTCTGTGTCTAGAACACGTGCTCATTCTTTCCCTAAAATGTATTCATAGG  
 TGAAAAAGGGCTTCTGCGGAAAGCCCTGGGTATGTGGGAAACCTGGATTTACAGCTGT  
 CTTTCCAGCAGGATGATGCAGGAGAGAGAGGGATGCGATTTCTCCCAATCTCTCTGGTC

16153 CGCCCCCTCAGCCCCGGCCCCGCCCTCCCTTCTGGCCCCGCCTCTGCCAGAGCCCTTCTC  
 AAGCCAGGAAAACCTGGTAATTTCTATTTGCCCTCTCCTCCTGTGGTTCTGCCCGGGCCCT  
 GAGGCGGGCTCTAAAGCCCTAGTCTCACCCCTCAAGAAGGAAGAAGTAGAGTCATCACCTC  
 TAAATCCCTCCTCCACACGGCCCCCTCCTCTATTGCAGATCCTGGGCATCTACAAGCAG  
 GGCTCAAATGCCAGGTGAGATGGAATGACTGGAAGGGCTGCTGGGCAGTGTTTTTTTT  
 [T, G]  
 TTTGTTTGTGTTTGGGAGAGTTACTATTTTGGTGGGGCAATTGCCAAGGAGTGAAGTA  
 CCTTAAATCAGAGGCGCATGGCCGGGCATGGTGGCTCAAGCCTGTAATCCCAGCACTTT  
 GGGAGGCCGAGGCGCGCAGATCACCTGAGGTGAGGTCAAGAGTCAAGACCAGCCTGACCAACAT  
 AGCGCAACCCCGCCTCTACTAAAAATACAAAAGTAGCTGGGCGTGGTGGCAGCCACCTG  
 TAATCCAGCTACTTGGGAGGCTGAGGCATGAGAATCGCTTGAACCTGGGAGGCGGGGT

FIGURE 3, page 10 of 12

16181 CTTCTGGCCCCGCTCTGCCAGAGCCCTTCTCAAGCCAGGAAAACCTGGTAATTCTATTT  
GCCTCTCCTCCTGTGGTTCTGCCCCGGGGCCCTGAGGCGGGCTCTAAAGCCCTAGTCTCAC  
CCTCAAGAAGGAAGAAGTAGAGTCATCACCTCTAAATCCCTCCACCACGGCCCTC  
CTCTATTGCAGATCCTGGGCATCTACAAGCAGGGCCTCAAATGCCGAGGTGAGATGGAAT  
GACTGGAAGGGCTGCTGGGCAGTGTCTTTTGTGTTGTTGTTGTTGTTGGGAGAGTTACT  
[G, A]  
TTTTGGTGGGGCAATTGCCAAGGAGTGAAGTACCTTAAAATCAGAGGCGCATGGCCGGGC  
ATGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCGGAGGCGCGCAGATCACCTGA  
GGTCAAGAGTTCAAGACCAGCCTGACCAACATAGCGCAACCCCGCTCTACTAAAAATAC  
AAAAAGTAGCTGGGCGTGGTGGCACCCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGC  
ATGAGAATCGCTTGAACCTGGGAGGCGGGGTTGCAGTGAGCCGAGATCACGCCACTGCA

16756 CAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGAGGGCTCTGTCTCA  
AAAAAAAAAAACAACAAAAAACCCCAAAACCAAAACCCACAAAATCAGAGGCTCAAG  
ATGACTGATGTGAAGGGAGTGGCGTTTAAAGAGGCCATTTATTTTGATGACGCAGCTGCCC  
AGGACAGAGAACATGGGAGAAGGCATAGACTGACAATTAGGAGGAGGAGAACACTTTGG  
AAGGAGACTCTTATTTTGGTGGGGCAGCTGCTCAGGAACAAAGGTTCTGGTAGGGGGC  
[A, G]  
CAAGCCTGCGGGATGGGATGGAGGGTATTCTGACCAATGTCCCTGGCTGGCTCTCCATTT  
GCTCTCCCCCAGCCTGTGGAGTGAAGTGCACAAAGCAGTGCAAGGATCGCCTGTGAGTTG  
AGTGTGCGGCGCAGGGCCAGAGTGTGAGCCTGGAGGGGTCTGCACCCTCACCTCACCCA  
TGCAAGCCACCATCACCGCGCCTTCAGCTTCTCTGCCCCGCTGGCAGGCGAGGCT  
CCAGGCCTCCAGGTAAGAGGGAGTCATTCTGTACTGGCCTGTGGAGGGAAGGATGCAGGG

18059 AATGATTATTTTGTCTGAGAACAGTCCGAACAACATATGTTAAACTGGGGTCTAAGGTAGTT  
GATCACAACCTGTTTGGGTGGGCATAAGTCCTCAAAAAACAGAGGCAGGCACAGGGCATAAC  
ATCCTCAAAAAATAGAAAAGATAAATCCATTGTCATTGAGCCTTCCAGAAGTGCTGGGGTCT  
TAAATGTGAAATACACAAAAATGACATTTAAGCAAACCTGCGCTGACAAATCTGTGGC  
TGAAAAAGCTGTGGCAAAACAAAAACATAGAAAAAGAGCCTCAAAAATTGGGCTGAGGCC  
[A, G]  
GGCATGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAAGCCAAGGTGGGTGGATCACC  
CGAGGTGAGGAGTTGGAGACCAGACTGGCCAACGTGGCAAAACATCATCTCTACAATACA  
AAAATACAAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGG  
CTGAGGCACGAGAATCGCTTGAACCTGGGAGGTGGAGGTTGCAGAGAGCCGAGATTGCGC  
CATTGCACTCCAGCCTGGGCGACAGAGAGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA

18364 TGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAAGCCAAGGTGGGTGGATCACCCGAG  
GTCAGGAGTTGGAGACCAGACTGGCCAACGTGGCAAAACATCATCTCTACAATACAAAAA  
TACAAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGA  
GGCAGGAGAATCGCTTGAACCTGGGAGGTGGAGGTTGCAGAGAGCCGAGATTGCGCCATT  
GCACTCCAGCCTGGGCGACAGAGAGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAA  
[A, -, T]  
TGGGCTGTGAGGTCATGCAGGGAATTGATTTTTTGGTGGGTGGGTCTGCTTCTGGGATGAT  
GTGGATGCCTCCCGTGGAGAGGGGAAGGGTTGATGAAGTCCCAGGGACCTGGAAGTGTGT  
TCTGCAGCAATCCCCCTCCCAGCAGAGATCCGTGAGGAGGAGGTACAGACGGTGGAGGAT  
GGGGTGTGTTGACATCCACTTGAATAGATGGTGAGTCTCCACAGCTGGCACCAGAGCT  
CCCCACTGAGGGCTGGGGGGGAGCTGGGGAGTATCAGGGAATGGGTGCTTTATCAAAT

18861 ACTTGTAATAGATGGTGAGTCCTCCACAGCTGGCACCAGAGCTCCCCACTGAGGGCTGG  
GGGGAGCTGGGGAGTATCAGGGAATGGGTGCTTTATCCAAATGGCTCCAAGCCAGGTG  
GGCTACTACCTTGTGTTAGGGGGGTGTCTTCTCACAACCTGTTTTTCTCTCCAGCT  
GTGGTTGGATCAAGGACTCATTCCTGCCTTGGAGAAAATACTTCAACCAGAGCAGGGAGC  
CTGGGGGTGTGCGGGCAGGAGGCTGGGGATGGGGGTGGGATATGAGGTGGCATGCAGCT  
[G, A]  
AGGGCAGGGCCAGGGCTGGTGTCCCTAAGGTTGTACAGACTCTGTGAATATTTGTATTT  
TCCAGATGGAATAAAAAGGCCCGTGAATTAACCTTCACCATCAGCGCCTAGAATCCCGG  
GGGGTAGGGGGATGTTATACCTTACAGGATGACAATCTTGGGAGCTAGAACTTTGTAGCC  
AGAGAACTTGGGAGGTCTGGAATCTCATGTGTCTGGAGTCTTGGGGAAGAGAATCTTAG  
AAGCAGAAAACCTTGAACATAAGAATCTTGGGAGGGTCTAGGATCTTGAGGAGACCAG

20443 GTTGTGAAGAGCAGCTCGCTCCTGTGCCGCTGCCTCCTGTGCTGCCTCCATCCCTGCA  
GCCAGTCGGTTCCTCTTGGCTCCTCTCGTCACTACCCTCCAGTTCAGTCTGGCCTCTT

FIGURE 3, page 11 of 12

CCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGCATGCATGCATATGTGTGTCC  
AGGTCTGCCTGCCCCGGGATGTGACAAGTAGCGGTCTTCATGGTTGCATGTGTCTGAATTT  
GGTGTCTGAGCTTCACATTGTATGCGCCTGTGTGCATGTGTGTGCATGGACATGCATGCT  
[G, A]  
TATCTGCTGTGTTTCCCCTCCCCCATGTGTCCCCTGGCCTTTGCACATGGGAGAAGGG  
CATGTGCTCAGCATATCACTCAACTGTCCACATTGGGTGGGTACCTGTGTGTGGTGTGTG  
TGTGTGGGGGGTGTGTCTTGAAGTGGCAGGTCCCAAATGCTTAGGCAATCTGAACCTTGG  
ACCTTGCAGAGAGGAGAGATGTCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGC  
CCGGTGACCTTTTCTGCCCTTGATGGGCAAAGCTGGGGGTAGGGAAAGGAGACAAGTGCT

20881

TTGAAGTGGCAGGTCCCAAATGCTTAGGCAATCTGAACCTTGGACCTTGCAGAGAGGAGA  
GATGTCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGCCCGGTGACCTTTTCTGC  
CCTTGATGGGCAAAGCTGGGGGTAGGGAAAGGAGACAAGTGCTCATACTTACCTCCCTCC  
CTGCCCAGGCTCCTCTGTAAGGGTCTGAGTCTGTCTGTGAGCCATTGCATCTGTCTGT  
CTATGCCCTGATGCCTGGATGGACAAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG  
[A, T]  
GTGTGAGGCTGCAGGAAGAGGAACAGTGGGGGATGGGCAGGAAAGTGGGCTGTGGGGTCA  
GGGAGGCGAT